

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 5, 2003, 23:24:25 ; Search time 4104 Seconds

(without alignments)  
6516.929 Million cell updates/sec

Title: US-09-914-168-2  
Perfect score: 4/27  
Sequence: 1 MSKPVLFANRSMFVALAAY.....TGVKESGNPKLHFPIGTGF 919

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2.1/USPRO/US09914168/runat\_28042003.151614.6787/app\_query.fasta.1.1095  
-DB=genemdb1 -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human4.0.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09914168 @CGN.1.1.2655 @runat.28042003.151614.6787 -NCPU=6 -ICPU=3  
-NO.XLPYX -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_hgt\_hum:\*  
31: em\_hgt\_inv:\*  
32: em\_hgt\_other:\*  
33: em\_hgt\_mus:\*  
34: em\_hgt\_pln:\*  
35: em\_hgt\_rod:\*  
36: em\_hgt\_mam:\*  
37: em\_hgt\_vrt:\*  
38: em\_sy:\*  
39: em\_hgtgo\_hum:\*  
40: em\_hgtgo\_mus:\*  
41: em\_hgtgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4727	100.0	2760	6	AX034847	AX034847 Sequence
2	4727	100.0	89047	6	AX067459	AX067459 Sequence
3	4581	96.9	2670	6	AX034849	AX034849 Sequence
4	617	13.1	11158	1	AE004682	AE004682 Pseudomon
5	604.5	12.8	1374	6	AX078576	AX078576 Sequence
6	532	11.3	10302	1	AE006218	AE006218 Pasteurel
7	494	10.5	10831	1	AE013668	AE013668 Yersinia
8	494	10.5	216050	1	AJ414157	AJ414157 Yersinia
9	486.5	10.3	24892	1	AE008906	AE008906 Salmonell
10	486.5	10.3	249050	1	AL627283	AL627283 Salmonell
11	454	9.6	14390	1	AE004323	AE004323 Vibrio ch
12	436.5	9.2	1737	6	AX033468	AX033468 Sequence
13	436.5	9.2	11833	1	U32752	U32752 Haemophilus
14	434.5	9.2	243184	1	AP002568	AP002568 Escherich
15	433.5	9.2	10819	1	AE000493	AE000493 Escherich
16	433.5	9.2	11760	1	AE005654	AE005654 Escherich
17	433.5	9.2	338534	1	EC00963	U14003 Escherichia
18	427.5	9.0	1731	6	AX033470	AX033470 Sequence
19	375.5	7.9	11870	1	AE012531	AE012531 Xanthomon
20	359.5	7.6	10839	1	AE012071	AE012071 Xanthomon
21	353.5	7.5	20310	1	AY010120	AY010120 Xanthomon
22	347	7.3	10074	1	AE003957	AE003957 Xylella f
23	319	6.7	944	6	AX473246	AX473246 Sequence
24	268	5.7	201050	1	AL646064	AL646064 Ralstonia
25	265.5	5.6	2379	6	AX155443	AX155443 Sequence
26	265.5	5.6	2379	6	AX202493	AX202493 Sequence
27	262.5	5.6	10029	1	AE010493	AE010493 Fusobacte
28	259.5	5.5	2379	1	NG081959	U81959 Neisseria g
29	259	5.5	2394	6	AX155439	AX155439 Sequence
30	259	5.5	2394	6	AX155447	AX155447 Sequence
31	259	5.5	2394	6	AX202489	AX202489 Sequence
32	259	5.5	2394	6	AX202497	AX202497 Sequence
33	259	5.5	13538	1	AE002375	AE002375 Neisseria
34	259	5.5	110000	2	AC020884.1	Continuation (2 of
35	259	5.5	340806	1	NMA122491	Continuation (2 of
36	259	5.5	349980	6	AX043922	AX043922 Sequence
37	255.5	5.4	6444	1	AF461795	AF461795 Bartonell
38	253	5.4	2394	1	AF461245	AF461245 Neisseria
39	245.5	5.2	181233	2	AC010535	AC010535 Homo sapi
40	244	5.2	10029	1	AE012236	AE012236 Xanthomon
41	242.5	5.1	189050	1	AL646069	AL646069 Ralstonia
42	241.5	5.1	10242	1	AE011773	AE011773 Xanthomon
43	240	5.1	333800	1	SME591792	AL591792 Sinorhizo
44	238.5	5.0	6617	6	AX139467	AX139467 Sequence
45	236.5	5.0	10029	1	AE008175	AE008175 Agrobacte

RESULT 1

## ALIGNMENTS

AX034847  
LOCUS AX034847 2760 bp DNA linear PAT 15-NOV-2000  
DEFINITION Sequence 1 from Patent WO052042.  
ACCESSION AX034847  
VERSION AX034847.1 GI:11190801  
KEYWORDS  
SOURCE Moraxella catarrhalis.  
ORGANISM Moraxella catarrhalis.  
Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
Moraxella.  
REFERENCE Ruelle, J. L.  
1 (bases 1 to 2760)  
AUTHORS Immunogenic compounds  
TITLE Patent: WO 0052042-A 1 08-SEP-2000;  
JOURNAL RUELLE JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE)  
FEATURES  
source  
1. 2760  
/organism="Moraxella catarrhalis"  
/db\_xref="taxon:480"  
BASE COUNT 775 a 345 c 654 g 786 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,68e-289 length: 2760  
Score: 4727.00 matches: 919  
Percent Similarity: 100.008 Conservation: 0  
Best Local Similarity: 100.008 Mismatches: 0  
Query Match: 100.008 Indels: 0  
DB: 6 Gaps: 0  
US-09-914-168-2 (1-919) x AX034847 (1-2760)  
OY 1 MetSerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaTyr 20  
DB 1 ATGCAAGCCCGTTTCTTGTGCAAAATCGAGTTTATGCTGCGCATGGCGCTTAT 60  
OY 21 LeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleLeu 40  
DB 61 TTGCGCTTGGATGACATCGCAAGCATTTGGCCACACAAATTAACCTCGCAACATCAT 120  
OY 41 HisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlnAsnProProValLeu 60  
DB 121 CATGTACCCCGCTCATGACACCGCATCAATCAGCAAAAGGCAAGCAATCCCGCTTTTG 180  
OY 61 LeuThrProGlnGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGln 80  
DB 181 CTAACACCTTGAGAGATACAAAGCAGCGCTTAATGCTGCTGAGCTGAATGCTAACGCCAA 240  
OY 81 SerGlnAlaLeuAspValIleAsnPheAspAspGlnSerProIleSerArgIleGlyGlu 100  
DB 241 TCACAGACCTTGGATGTTGTTCAATTTTGATGATCAATCGCGATATCTCGATCGGTAG 300  
OY 101 GlnSerProProLeuGlyLeuAspMetSerValIleGlnGlnIleThrThrProLeuSerLeu 120  
DB 301 CAATCACCCCGCTTGGGTTGGATGATGTCGTCATCGAAGAAACACACCGCTAAAGCTTG 360  
OY 121 GlnGluLeuPheAlaGlnGlnIleSerThrGlnMetGlyIleAsnProAsnAspTyrIlePro 140  
DB 361 GAGCAATTAATTTGCTCAAGATCTACTGAGATGGGAATCAATCCAAATGATTATATTC 420  
OY 141 GlnTyrGlnGlnGlnGlnIleProAsnSerGlnValAlaValProProThrLeuGlnProGln 160  
DB 421 GAATATTCAGGCGCAGCAACCTTAATAGTAGGTGGTTGATACACCGACCTTAAGACCTGAA 480  
OY 161 LysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysVal 180  
DB 481 AAACACGAGTTTGATCAAGCGCTTATGACAGCGCTATTTATGATGGTGCAATAAGGTG 540  
OY 181 ProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGlnThrSerAlaIleGly 200  
DB 541 CTTAAGCCTTAAGGCAAAATTTATCAATCATCGCAATCAGCGAAGACAGCTGCGATTGGG 600  
OY 201 SerSerIleGlnLysThrGlnProTyrAlaAsnIleLysAlaIleLeuGluAspIleThr 220

DB 601 TCATCGCATCAAAAAACAGACCTTATGCAATATCAAAAGCAGCATGAGACATCAC 660  
OY 221 GlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgIleThrAlaLeuVal 240  
DB 661 CAGAGTACGAGTGGATGATTAAGGCTTATCCACCGCTTAAGGCAAAACGCTTGGTG 720  
OY 241 AlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGly 260  
DB 721 GCACGCGCTGCTGTCGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
OY 261 GluValAspValIleIleHisAspLeuGlyLysProValTyrIleAspTyrArgAlaVal 280  
DB 781 GAGGTGATGATCATCATCCATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
OY 281 GluValArgGlyGlnGlnAlaAspAspLysAlaPheThrThrValAlaAspGluValPro 300  
DB 841 GAGGTGAGGTGAAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
OY 301 LeuLeuIleGlyAspValPheHisHisGlyLysTyrGlnThrLysAsnLeuIleGlu 320  
DB 901 TTGCTGATCGCGCATGCTCTTTCATCATGCAAGTACGAACCAAAATCTCATCGAA 960  
OY 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrPheAspAspArgSerValAspVal 340  
DB 961 AATGCCAGTCTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
OY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlyTyrArg 360  
DB 1021 ATTTTCCAGATTAATACCGCTGATGCTGCTTAATTTTGTATACAGTACCATATGCT 1080  
OY 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380  
DB 1081 TTTGATGAGGTGATATTTTATTCATTTGATCTTAACCAATCAATGCAACCGATCCA 1140  
OY 381 AspLysLeuProValLysArgGluLeuLeuGlnGlnLeuThrValAsnMetGlyGlu 400  
DB 1141 GATTAAGCTGCCAGTTAAACGGAATTAATTAAGCTTAAGCTTAAGCTTAAGCTTAAGCT 1200  
OY 401 AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420  
DB 1201 GCTTACATTTTACAGGCGGTGCTGACTTCAATATGATTTGATTTGCCACAGGATTTT 1260  
OY 421 AsnMetValAsnThrGlnIleValPheProGlnArgGlnGlnIleGlnAsnAspGlnVal 440  
DB 1261 AATATGGAATTAACGGAATTTGCTTCCAGAGGTGAACGATCAAAAGCACCAATG 1320  
OY 441 SerPheGlnGlnSerSerSerSerArgThrGlnProAlaGlnValAspGlnSerThrLeu 460  
DB 1321 AGCTTTGAGCGTCTTCAAGTAGCCGTAAGTACACAGCACAAAGTTGAGAAAGCACCT 1380  
OY 461 GluProValIleGlnThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIle 480  
DB 1381 GAACCTGCTATGAACCGTTGAGCTAAGCATGAGATTAATTAAGATTTTCCGCCATC 1440  
OY 481 GluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg 500  
DB 1441 GAATTTAGTGCATCTAATCTCATTTCAAGACCAAGCTTAATTTGGCGCTGCCAAGGCTCG 1500  
OY 501 HisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
DB 1501 CATTTATATGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
OY 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540  
DB 1561 CGCTCTATTTTGGGAGAGATGAGGATGCGCTATCTGCGCTGACGCTATTTTACT 1620  
OY 541 AspGlnSerGluAsnGluValIleAspLeuProGlnArgThrAlaLeuAlaAsnArgLys 560  
DB 1621 GATGAATCTGAATAATGAGGTAAATGATTTGCCAGAGCTACCGCATTTGCTAAATCGCAAG 1680  
OY 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValAlaPheValAlaSerAsp 580

Dh 1681 ACCCTGCTGATGCTATATCAAGTAAAAAGTCCGCTATATGCTTTGCGGAGTGAT 1740  
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600  
Dh 1741 AACCACGAGATGTCGCAATTTGTTGGCTGGGATCGGACACAGATACCCGCTAGTC 1800  
Qy 601 ThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrglnAlaGlyAlaGluLeuArg 620  
Dh 1801 ACAAAATTTGACGCAATATGATTAATCGTATGATGCTATCAAGACGCGCTAGCTAAGA 1860  
Qy 621 LeuSerGluAspLysLysGlyValLysLeuTyrglnAlaThrLysProLeuSerHisProLeu 640  
Dh 1861 CTGTCTGAGGATAAAAAAGGGGTCAAGTTATATGCCCAACCCCTTAGCCACCTCTTA 1920  
Qy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrglnGlnGlnValPheGlyHisSerThrAsn 660  
Dh 1921 AATGATCACTAGACGCAACTTTGGGTTATCAACAAGAGTTTGGTCACTCTACCAAT 1980  
Qy 661 GlyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGlnAsn 680  
Dh 1981 GGTTTGTATTTATCCACACGACCGCTAGAGCATGAGATTAGCCGAGTATTATCCAAAT 2040  
Qy 681 GlyGlyTrpAsnArgThrThrSerLeuArgTyrglnLeuAspLysLeuLysThrGlnAla 700  
Dh 2041 GGTGGCTGGAATCGTACTTATTCATTCGTTATCGCTTGATTAACCTTAAACCCAGCA 2100  
Qy 701 ProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlu 720  
Dh 2101 CCCCCGTAACATGCGAGGATTTACCAGTGGATTTTGCTCAATGCTAAGCCAAAGCCCAAGG 2160  
Qy 721 AlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet 740  
Dh 2161 GCGTATTATGGCAGGTGCTGCTGCATAAACGTTGCGATTAATTTGCTTAATCCGATG 2220  
Qy 741 ArgGlyTrpArgGlnArgTyrglnSerLeuGlnValGlySerSerGlyLeuValSerAspAla 760  
Dh 2221 CGTGGCTATCGTCACGAGATTTCTTTAGAGTTGGCTCAAGCGTTTGGTATCGGATGCT 2280  
Qy 761 AsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyrglnPheGlyAspAsnAlaTyrg 780  
Dh 2281 AATATGGCATATTCGTCGAGCTGATTAAGTGGCGGTATAGTTTGGGATATATGCTTAT 2340  
Qy 781 GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrglnLeuProSerAsp 800  
Dh 2341 GGCACCAATCGTCCCATCAGATGACTGGTGGCATCAACAGCAGATACATTGGTCGAGT 2400  
Qy 801 AsnPheAsnHisValProTyrglnLeuArgPhePheAlaGlyLysAspGlnSerIleArg 820  
Dh 2401 AATTTTAATCATGTCCCATATCGTTGGCTTTTGGCTGGTGGCAGCCAAAGTATTCGT 2460  
Qy 821 GlyTyrglnAlaHisAspSerLeuSerProIleSerAspLysGlyTyrglnLeuThrGlyGln 840  
Dh 2461 GGAATATGCACATGACAGATTTATACCTATATCAGATTAAGGTATATCTGACAGCGGTCA 2520  
Qy 841 ValLeuAlaValaGlyThrAlaGlnTyrglnAsnTyrglnPheMetLysAspLeuArgLeuAla 860  
Dh 2521 GATATGGCGGTGTACAGCTGAATATATATGAATTATGAAGATTTCGCTTGGCGC 2580  
Qy 861 ValPheGlyAspIleGlyAsnAlaTyrglnAspLysGlyPheThrAsnAspThrLysIleGly 880  
Dh 2581 GTTTTGGTATATGATTAATGCTTATGAATAAGGCTTTACTAAATGATACCAAAATTTGGT 2640  
Qy 881 AlaGlyValaGlyValaArgTrpAlaSerProValaGlyGlnValaArgValaAspValaAlaThr 900  
Dh 2641 GCAGGTGTCGCTGCTCGCTGGCATCACCTGTCGCTCAAGTTCTGCTTATGTCGCACT 2700  
Qy 901 GlyValaLysGlnGlyLysAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919  
Dh 2701 GGTGTCAAAAGAGAGGCAATCCATTAAGTCATTTTATATGTCACACCATTT 2757  
RESULT 2  
AX067459  
LOCUS AX067459 89047 bp DNA linear PAT 24-JAN-2001

DEFINITION Sequence 34 from Patent WO0078968.  
ACCESSION AX067459  
VERSION AX067459.1 GI:12545079  
KEYWORDS  
SOURCE Moraxella catarrhalis.  
ORGANISM Moraxella catarrhalis  
REFERENCE 1 (bases 1 to 89047)  
AUTHORS Lagace,R.E., Patterson,C. and Berg,K.L.  
TITLE Nucleotide sequences of moraxella catarrhalis genome  
JOURNAL Patent: WO 0078968-A 34 28-DEC-2000;  
Incyte Genomics, Inc. (US)  
FEATURES  
source location/Qualifiers  
1..89047  
/organism="Moraxella catarrhalis"  
/db\_xref="taxon:480"  
unsure  
BASE COUNT 26501 a 17338 c 19060 g 26147 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,56e-287 Length: 89047  
Score: 4727.00 Matches: 919  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-914-168-2 (1-919) x AX067459 (1-89047)  
Qy 1 MetSerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrg 20  
Dh 80397 ATGTAAACCCCGCTTTGTTGTTCCAAATCGCACTTTATGCTGTGCGCATTTGCGGCTTAT 80456  
Qy 21 LeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsn 40  
Dh 80457 TTGCCCTTGATGATGACATCGCAACCATTTGGCGCAACAATAATACCTGCAACATCATCAT 80516  
Qy 41 HisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyLysAsnProProValLeu 60  
Dh 80517 CATGTACCCGCTCAAGACACCCCATCAATCAAGCAAAAGCGAGCAATCCGCTGTTTGG 80576  
Qy 61 LeuThrProGlnGlnIleGlnAlaArgLeuAsnAlaAlaGlyLysAsnAlaLysProGln 80  
Dh 80577 CTAACACCTGAGACGACATACAGACCGCTTAATGCTGCTGGAGCTGAATGCTAAAGCCCCAA 80636  
Qy 81 SerGlnAlaLeuAspValaValaAsnPheAspGlnSerProIleSerArgIleGlyGln 100  
Dh 80637 TCACAAGCTTTGGATGTTGTCAATTTTGTATGATCAATGCCGATATCTCGTATCGGTGAG 80696  
Qy 101 GlnSerProProLeuGlyLeuAspMetSerValIleGluGlnThrThrProLeuSerLeu 120  
Dh 80697 CAATCACCCCTTTGGGTTGATATGTCGTCATCGAAGAAACACACCGCTAAGCTTG 80756  
Qy 121 GlnGlnLeuPheAlaGlnGlnSerThrGlnMetGlyIleAsnProAsnAspTyrglnPro 140  
Dh 80757 GAGGATATTTTGGCTCAAGAACTCTAGATGAGGAAATCAATCAAAATGATTAATTTCCA 80816  
Qy 141 GlyTyrglnGlyGlnGlnProAsnSerGlnValaValaProProThrLeuGlnProGln 160  
Dh 80817 GAATATCAAGGCGACACCACTTAATAGTGAAGTGTTGATACCGCAATTAACACTGAA 80876  
Qy 161 LysProGlyLeuIleLysArgLeuTyrglnAlaArgLeuPheAsnAspGlyValaLysLysVal 180  
Dh 80877 AAACCAAGTTTGATCAAGCGCTTTATATGCACCCCTTAATTAATGATGCTCATTAAGGTG 80936  
Qy 181 ProArgLeuLysAlaLysPheTyrglnSerSerGlnSerGlnGlyGlnThrSerAlaIleGly 200  
Dh 80937 CCTAGGCTTAAAGGCAAAATTTTATATCATATCCATCAAGCGGAACCAAGTGGCTTTGGG 80996  
Qy 201 SerSerHisGlnLysThrGlnProTyrglnAlaAsnIleLysAlaAlaLeuGlnLysPheThr 220

Db	80997	TCATCCCATCAAAAAACAGACGCTTATGCAAAATATCAAAAGCAGCACTTAACACATCAC	81056
Oy	221	GLNGLSerAlaMetAspLeuAsnGlySerIleProArgLeuAlaArgGlnThrAlaLeuVal	240
Db	81057	CAAGATCGACGCATGATTTGAATGAGCTCTATCCACGCCCTAAGCAACTCTTTGGTG	81116
Oy	241	AlaAlaArgAlaValAlcIlyrThrAspIleAspLeuSerIleIlaArgAsnSerIleGly	260
Db	81117	GCAGCCCGGCGCTCGGTATTTATGATATTGATTATCATATCATATGAAGAAATGCAATCGGA	81176
Oy	261	GIuValAspValIleIleHisAspLeuGlyGIuProValTyrIleAspTyrArgAlaVal	280
Db	81177	GAGGTGATGTCATCATCATCATGATTAGTGAACCGTTTATATGATTATGAGAGGGTG	81236
Oy	281	GIuValArgGlyGIuGIuValAspAspIlyAlaPheThrThrValAlaAspGIuValPro	300
Db	81237	GAGGTACGAGGGAAGCTGCTCATGATGAAGCATTTACACCGCTGGCGATAGGGTCA	81296
Oy	301	LeuLeuIleGIuAspValPheHisHisGlyLysTyrGIuThrLysLysAsnLeuIleGIu	320
Db	81297	TTGCTGATCGCGGATGCTCTTTCATCATCGCAAGTACGAACCAAAAAATCTCATCGAA	81356
Oy	321	AsnAlaSerAlaGIuHisGlyTyrPheAspGIuArgTyrLeuAspArgSerValAspVal	340
Db	81357	AATGCCAGTGCAGAACATGATATTGTTGAGGGCGGTGGCTGATCGTTCATGTATGTA	81416
Oy	341	IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg	360
Db	81417	ATTTTCCAGATATATACCGCTGATGTCAGCTTAATTTATGATACAGGTACGCAGTATCC	81476
Oy	361	PheAspGIuValAlaPhePheThrIleAspProLysTyrAsnGlnLeuThrThrAspPro	380
Db	81477	TTTGATGAGAGTGGTATTTTTCACATGATTCCTAAACCAATCAATTCACACCGATCCA	81536
Oy	381	AspLysLeuProValLysArgGluLeuLeuGlnGlnLeuLeuThrValAlaAsnMetGIuGIu	400
Db	81537	GATTAACCTGCCAGTTAACAGAGAAATTACTTGAACACTTACTACCGCTTAACATGGGAGG	81596
Oy	401	AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe	420
Db	81597	GCTTACAATTTACAGCGCGGTGCGTCACTTCAAAATGATTGATGGCCACACGATATTTT	81656
Oy	421	AsnMetValAsnThrGIuIleValPheProGluArgGIuGlnIleGlnAsnAspGlnVal	440
Db	81657	AATATCGTAAATACCGAGATGTCCTTCCAGAGCGGAACACATCCAAACACCAAGTG	81716
Oy	441	SerPheGlnGlySerSerSerSerArgTyrGlnProAlaGlnValAspGIuSerThrLeu	460
Db	81717	AGCTTGACAGAGCTTCAGTAGGCGCTACTGAACACGACCAAGTTGATGAAGCACACTT	81776
Oy	461	GIuProValIleGIuThrValGIuLeuThrAspGlyIleLeuMetAspIleSerProIle	480
Db	81777	GAACCGTGTATTTCAAAACCGTTGAGCTAACGAGATGGGATATTATGATATTTTCGCCCATC	81836
Oy	481	GIuPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg	500
Db	81837	GAATTTAGTGCATTAATCTGATTCAAGCAAGCTAAATTTGTGTGCTCCCAAGCTCCG	81896
Oy	501	HisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGIuValAsn	520
Db	81897	CATTATATGACATGCCGTGATGATAGGGTGTCCATCAATCATGATGATGGCGTAAT	81956
Oy	521	ArgSerIleLeuGIuTyrArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro	540
Db	81957	CCCTCATTTTTGGCGAATCAGCGGTGCGGATCTCGCGGTGCACAGTCTATTTTTACT	82016
Oy	541	AspGIuSerGIuAsnGIuValIleAspLeuProGIuArgThrAlaLeuAlaAsnArgLys	560
Db	82017	CATGATATCGAAATAGGTAATAGATTGGCCGACGCGTACCGCATGGCTATATCGCAAG	82076
Oy	561	ThrProAlaAspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAsp	580

Db	82077	ACCCTGCTGATGTCGTATCCAAAGTAAAGTGGCCGATATGTCCTTTGTGGCGAGTAT	82136
Qy	581	lysProArgaspGlyGlnIleGlyLeuGIyTrpGlySeraspThrGlyThrArgLeuVal	600
Db	82137	AAACACGACGATGTCGCAAAATTGGTTGGGCGGGGATGCGACACAGGTAACCCGGCTAGTC	82196
Qy	601	ThrLysPheGluHisAsnLeuIleAsnArgaspGlyGlyTrpGlnIleGlyValGlyLeuArg	620
Db	82197	ACAAATTTTGACATAAATTTGATTATGTCGTATGGCTATCAAGGCGGGCTGAGCTAAGA	82256
Qy	621	LeuSerGluaspLysLysGlyValLysLeuTyraLthrLysProLeuSerHisProLeu	640
Db	82257	CTGTCTGAGGATAAAAAAGGGGTCAAGTTATATGCCACCAACCGCTTAGCCACCGCTCA	82316
Qy	641	AsnaspGlnLeuArgAlaThrLeuGlyTrpGlnGlnGluValPheGlyHisSerThrAsn	660
Db	82317	AATGATACGCTAAGAGCAACTTTGGGTTATGACACAGAAAGTTTGGTCACGTACCAAT	82376
Qy	661	GlyPheaspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsn	680
Db	82377	GGTTTTGATTTATCCACACGCAACCTTAAGCATGATGATTACCCGACGATTTATCCAAAT	82436
Qy	681	GlyGlyTrpAsnArgThrTrpSerLeuArgTyraArgLeuaspLysLysThrGlnAla	700
Db	82437	GGTGCTCGAATTCGACTTATTCATTGCGTTATCGCTTATCGCTTAAGCTTAAACCCAGCA	82496
Qy	701	ProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlu	720
Db	82497	CCCCGTGAACATGGCAGAGATTTACCAGTGGATTTTGTCATATGGTAAGCCAAACCCAAAG	82556
Qy	721	AlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet	740
Db	82557	GCGTTATTGGCGAGCGTGTGCTGCTGATAAACCGGTGCAGATTAATTGGTTAAATCCGATG	82616
Qy	741	ArgGlyTrpArgGlnArgTrpSerLeuGlnValAlaGlySerSerGlyLeuValSerAspAla	760
Db	82617	CTGGCTATCGTCACGCGATTTCTTTACAGTTGGCTCAACGCGTTTGGTATCCGATGCT	82676
Qy	761	AsnMetAlaIleAlaArgAlaGlyLleSerGlyValTyTrSerPheGlyAspAsnAlaTrp	780
Db	82677	AATATGGCTATATGCTCGAGCGTGTATGTAGTGGCGGTATAGTTTGGGGATAAATGCTTAT	82736
Qy	781	GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTrpIleTrpSerAsp	800
Db	82737	GGCACACATCGTCCCATCAGTACTGTGTGGCATACAGCAGGATACATTTGGCGGAT	82796
Qy	801	AsnPheAsnHisValProTyraArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg	820
Db	82797	AATTTTAAATCATGTCGCAATATCGTTTGGGTTTTTTCGTGGGGGACCAAAATATTGCT	82856
Qy	821	GlyTyraIleHisAspSerLeuSerProIleSerAspLysGlyTyraLeuThrGlyGlyGln	840
Db	82857	GGATATGCACATACAGAGTTTATCACCTTATATCAGATAAGGGTTATCTACACAGGCGCTAA	82916
Qy	841	ValLeuAlaValaGlyThrAlaGlyTyraAsnTyraGluPheMetLysAspLeuArgLeuAla	860
Db	82917	GTATTGGCGCGTTGGTACACGCTGAATATATATATGATTTATGCAAGATTTTCGCTTTGGCG	82976
Qy	861	ValPheGlyAspIleGlyAsnAlaTyraLysAspLysGlyPheThrAsnAspThrLysLleGly	880
Db	82977	GTTTTTGGTATATGTAATGCTTATATATAAGGCTTTACTAATGATNCCAAAAATTGGT	83036
Qy	881	AlaGlyValaGlyValaArgTrpAlaSerProValaGlyGlnValaArgValaAspValAlaThr	900
Db	83037	GAGAGTGTGGGTTCGCTGGGCAATCACCTGCGGTCAAGTTCGTTGATGTGGCAACT	83096
Qy	901	GlyValaLysGlnGlnGlyAsnProLleLysLeuHisPhePheIleGlyThrProPhe	919
Db	83097	GGTGTCAAAAGAAAGAGGCAATCCCATTTAAGCGTCATTTTTTTTATTTGCGACACACATTT	83153
RESULT 3			
AN034849			
LOCUS			
AX034849			
2670 bp			
DNA			
linear			
PAT 15-NOV-2000			



Tue May 6 09:37:16 2003

us-09-914-168-2.rge

age 5

Sequence 3 from Patent WO0052042.  
ACCESSION AX034849 GI:11190802  
DEFINITION Moraxella catarrhalis.  
KEYWORDS Moraxella catarrhalis.  
SOURCE Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
ORGANISM Moraxella.  
REFERENCE 1 (bases 1 to 2670)  
AUTHORS Ruelle, J. L.  
TITLE Immunogenic compounds  
JOURNAL Patent: WO 0052042-A 3 08-SEP-2000;  
RUELLE JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE)  
FEATURES  
source location/Qualifiers  
1.2670  
/organism="Moraxella catarrhalis"  
/db\_xref="taxon:480"  
BASE COUNT 757 a 527 c 632 g 754 t  
ORIGIN  
Alignment Scores: 8,72e-280 Length: 2670  
Pred. No.: 4581.00 Matches: 888  
Score: 100.00% Conservative: 1  
Percent Similarity: 99.89% Mismatches: 0  
Best Local Similarity: 96.91% Indels: 0  
Query Match: 6 Gaps: 0  
US-09-914-168-2 (1-919) x AX034849 (1-2670)  
QY 31 GlnGlnAsnProAlaAsnTlleLleAsnHisValProAlaHisAspThrAlaIleAsn 50  
Db 1 CAACAATTAACCTCGAACAATCATCATCATCTACCGCTCATGACACCGCATCAAT 60  
QY 51 GlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnIleGlnAlaArgLeu 70  
Db 61 CAAAGCAAGCGAGCAATCCCGCTGTTTGCTACACCTGAGCAGTACAGACGCCCTT 120  
QY 71 AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValIleAsnPha 90  
Db 121 AATCGCTGCTGACTGAATGCTTAAGCCCAATCAACAGCTTGATGCTCAATTTTGA 180  
QY 181 GATCAATCGCGATATCTGATCGTATCGGTGACATACCCCTTGAGTTGATATGTCG 240  
Db 91 AspGlnSerProIleSerArgIleGlyGlnIleAsnProLeuGlnIleAsnPha 110  
QY 111 ValIleGlnIleThrProLeuSerLeuGlnIleAsnPhaIleGlnIleAsnPha 130  
Db 181 GATCAATCGCGATATCTGATCGTATCGGTGACATACCCCTTGAGTTGATATGTCG 240  
QY 241 GTATCGAAGAAACACACCGCTGATGCTGAGATTTGCTCAAGATCTACTAGAG 300  
Db 111 ValIleGlnIleThrProLeuSerLeuGlnIleAsnPhaIleGlnIleAsnPha 130  
QY 301 ATGGGAATCAATCAATGATTTATTCAGATATTCAGAGCAGACCTTAATGAGAG 360  
Db 151 ValIleValProProThrLeuGlnIleAsnPhaIleGlnIleAsnPha 170  
QY 361 GTGGTTGTACACGACATTTGAACTGAAACCCAGGTTTGATCAAGCCCTTTATGCA 420  
Db 171 ArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheArg 190  
QY 421 CCCATTATTATGATGCTGATCAATAGTGTGCTTGAAGCAAAATTTTATCATCA 480  
Db 191 SerGlnSerGlyIleThrSerAlaIleGlySerSerHisGlnIleGlnIleAsn 210  
QY 481 TCCCATTCAGCGAAGACAGTGCATGCGCTATCGCATCAAAACAGAGCCTTATGCA 540  
Db 211 AsnIleLysAlaIleGlnIleAsnIleGlnIleAsnIleGlnIleAsnIleGln 230  
QY 541 AATATCAAGAGCACTGTAAGACATCAACCAAGTCAAGCGATTTGATGAGCTCT 600  
Db 231 IleProArgLeuArgGlnIleThrAlaLeuValAlaIleArgAlaValGlyTyr 250  
QY 601 ATCCACGCGCTTAAGCAAACTGCTTGTGGAGCGCGCTGCTGCTGATTTATGATAT 660

QY 251 AspleuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGly 270  
Db 661 GATTATCAATCAATTAAGAAATAGCATCGAGAGCTGATCATCATCATCATGATTAAGT 720  
QY 271 GlnProValTyrIleAspTyrArgAlaValGluValArgIleGlyValAspAspLys 290  
Db 721 GAACCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 291 AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGly 310  
Db 781 GCATTACTACCGTGGCGGATGAGTGCCTTCTGATCGCGATGCTTCTCATCAGCGC 840  
QY 311 LysTyrGlnThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAsp 330  
Db 841 AGATCGAAGCAAAACAAATCTCATCGAATTCAGATGCTGCAATGATGATTTGAT 900  
QY 331 GlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSer 350  
Db 901 GGGCGTGGCTGATCGCTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 351 LeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAsp 370  
Db 961 TTAATTATGATACAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
QY 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeu 410  
Db 1021 CCTAAACCAATCAATTAAGCAACCGATCCAGTACGATGATGATGATGATGATGATGAT 1140  
QY 391 GlnGlnLeuLeuThrValAsnMetGlyValAlaTyrAsnLeuGlnAlaValArgAlaLeu 430  
Db 1081 GAGCAGTACTCCGCTTACATGATGAGAGCGCTTACAAATTTACAGCGCTGCGCATCT 1200  
QY 411 SerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhePro 450  
Db 1141 TCAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
QY 431 GlnArgGlnGlnIleGlnAsnAspGlnValSerPheGlnIleSerSerSerArgThr 470  
Db 1201 GAGCGTGAACATCAATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1320  
QY 451 GlnProAlaGlnValAspGlnSerThrLeuGlnProValIleGlnThrValGlnLeuThr 490  
Db 1261 GAACCGACACAGTATGATGAAGCAACCTGATGATGATGATGATGATGATGATGATGAT 1380  
QY 471 AspGlyIleLeuMetAspIleSerProIleGlnPheSerAlaSerAsnLeuIleGlnAsp 510  
Db 1321 GATGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
QY 491 LysLeuAsnLeuValAlaIleLysAlaArgHisLeuTyrAspMetProAspArgVal 530  
Db 1381 AAGCTAAATTTGGTGGCTGCCAAGCGCTGCCATTTTATGATGATGATGATGATGATGAT 1500  
QY 511 LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlnIleArgIleSerAspAla 550  
Db 1441 CTTGCCATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
QY 531 ValSerAlaValAlaArgAlaIleLeuProAspGlnSerGlnAsnGlnValIleAspLeu 570  
Db 1501 GTATCTCGCGCTGACGCTGATTTTACCTGATGATGATGATGATGATGATGATGATGAT 1620  
QY 551 ProGlnArgThrAlaLeuAsnArgLysThrProAlaAspValTyrGlnSerLysLys 590  
Db 1561 CCCAGCGCTACCGATTTGGCTATGCAAGACCCCTGCTGATGATGATGATGATGATGATGAT 1680  
QY 571 ValProLeuTyrValPheValAlaSerAspLysProArgAspGlnIleGlyLeuGly 610  
Db 1621 GTGGCGCTGATGCTTTGTCGCGAGTGAAGACCCCGAGATGCTCAAAATTTGTTGGCGC 1740  
QY 591 TrpLeuSerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 650  
Db 1681 TGGGATGCGACACAGTACCGCTTGTGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

OY	611	ASPGIYTYRGINALAGLYALAGLULEUARGLEUSERGILUASPLVSLYGLYVALLYSLEU	630
DB	1741	GATGGCTATCAACGACGGCGGTGCGTAAGACTGCTTGAGGATTAATAAAGGGGTCAAGTTA	1800
OY	631	TYRATATPRLVSPROLESENTHISPROLEUSNAASPGINLEARGALATHTHLEUGLYTYR	650
DB	1801	TATGCCACCAAAACCCCTTAGCCACCCTCTTAATGATGTCAGCTTAAGAGCAACTTTGGGTTAT	1860
OY	651	GLINGLIGLUALPHEGLYHISSETHRASNGLYPHEASPLEUSERTHYRGTHTHLEUGLU	670
DB	1861	CACAAAGAAGTTTTGGTCACTCTACCAATGGTTTGGATTTATATCCACACGACACCCTAG	1920
OY	671	HISGLUIIISERARGSETTLEILEGLNANGLYGTTPRANARGTHYRYSERLEUARG	690
DB	1921	CATGAGATTATGACCGGAGATTAATCCAAAATGGGTGGCTGGAATCGATATTATTCATTTGGT	1980
OY	691	TYRATGLEUASPLVGLLEULYSTRGCLINALAPROPFGLUPTHPRGLASPLEUPROYAL	710
DB	1981	TATCGTCTTGATTAACCTTAAACCCACACACCCTGAAACATGGCAGAGATTTACCAATG	2040
OY	711	ASPPHEVALASNGLYLVSPROSERGLNULALEULEUALAGLYVALALAVAHISLYS	730
DB	2041	CATTTTGTCAATGCGTAACCCAAAGCCAAAGCGGTATTGGCAGGCTGTGTCGATGATAA	2100
OY	731	THRVALLIASPASNLEUVALASPRKMEIARNGIYTYRARGINARGTYSERLEUULU	750
DB	2101	ACGGTTGACAGATAATTTGGTTAATCCGATCCGTGGCTATGTCACCGCATATTCTTTAG	2160
OY	751	VALGLYSERSEGLYLEUVALSERASPAIASMECLALILEALARGALAGLYLESER	770
DB	2161	CTTGGCTCAAGCGGTTTGGTATCGGATGCTAATATAGCGTATTGCTCGAGCTGGTATTAGT	2220
OY	771	GLYVALTYRSERPHEGLYASPAASNALATYRGILYSERASNARGALANHISGLIMETHRGLY	790
DB	2221	GCGCTGTATAGCTTTGGGGATATATCTTAATGCGACCAATCGGCCATCAGATGACTGGT	2280
OY	791	GLYILEGLINALAGLYTYRILETPSERSPASNPHEDASNHSVALPROTYRARGLEUARG	810
DB	2281	GCCATACCAAGCAGATACATTTGGTGGATAATTTTAAATCATGTGCCATATCGTTGGCT	2340
OY	811	PHENEAALAGLYLASPCINSERTLEARGLYTYRALANHSASPSERLEUSERPROLE	830
DB	2341	TTTTTTCCTGGCGGGACCAAGATTTGCTGGATATGCAATGACATGACAGTTTATCACTATA	2400
OY	831	SERASPLVSGLYTYRLEUETHRGLYGLYGLINVALLEUALAVAGLYTHRALAGLUYRASN	850
DB	2401	TCAGATTAAGGGGTATCTGACACGCGGTCAAGATTAATGGCGGTGGTAAGCCTGAATATAT	2460
OY	851	TYRGLUPHEMELLYASPLEUARGLEUALAVALEPHEGLYASPILEGLYASNALATYRASP	870
DB	2461	TATGAATATTATGCAAGATTTCGCTTGGCGGTTTTTGGTGAATATGGTAATGCTTATATAT	2520
OY	871	LYSGIYPIETHRASASPETHRSLYSILEGLYALAGLYVALIGLYVALRTPRALASERPRO	890
DB	2521	AAAGCCTTTACTAATGATATACCAAAATTTGGTGCAGGTGTCGGGTGTCGGTGGCCATCACC	2580
OY	891	VALIGLYGLINVALARGVALASPYVALATHTHRGLYVALLYSGIUGLUGLYASPROLIEUS	910
DB	2581	GTCCGTCAAGATTGCGTTGGATGTGCGCACTGCTGTCAAAAGAGAGGCAATCCATTAAAG	2640
OY	911	LEUHLSPHEPHEITLEGLYTHRPROPHE	919
DB	2641	CTGCATTTTTTTATTTGCGCACACCATT	2667
RESULT 4			
LOCUS	AE004682.C	1158 bp	DNA linear BCT 30-AUG-2000
DEFINITION	Pseudomonas aeruginosa PA01, section 243 of 529 of the complete genome.		
ACCESSION	AE004682	AE004091	
VERSION	AE004682.1	GI:9948598	
KEYWORDS	Pseudomonas aeruginosa.		
SOURCE			

ORGANISM	REFERENCE	TITLE	JOURNAL MEDLINE	REFERENCE	TITLE	JOURNAL	FEATURES
Pseudomonas aeruginosa	1 (bases 1 to 11158)		20437337				
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.	1 (bases 1 to 11158)		10984043				
Stover, C. K., Pham, X. O., Erwin, A. L., Miziouchi, S. D., Warrenner, P., Hickey, M. J., Brinkman, F. S., Huftagle, W. O., Kowalik, D. J., Lagrou, M., Garber, R. L., Goltzy, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L. L., Coulter, S. N., Folger, K. R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G. K., Wu, Z., and Paulsen, I. T.	2 (bases 1 to 11158)		10984043				
Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen	2 (bases 1 to 11158)		10984043				
Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA	2 (bases 1 to 11158)		10984043				
Location/Qualifiers	2 (bases 1 to 11158)		10984043				
1. 11158	2 (bases 1 to 11158)		10984043				
/organism="Pseudomonas aeruginosa"	2 (bases 1 to 11158)		10984043				
/strain="PAO1"	2 (bases 1 to 11158)		10984043				
/db_xref="taxon:287"	2 (bases 1 to 11158)		10984043				
complement(76..711)	2 (bases 1 to 11158)		10984043				
/gene="PA2541"	2 (bases 1 to 11158)		10984043				
complement(76..711)	2 (bases 1 to 11158)		10984043				
/gene="PA2541"	2 (bases 1 to 11158)		10984043				
/codon_start=1	2 (bases 1 to 11158)		10984043				
/transl_table=11	2 (bases 1 to 11158)		10984043				
/product="probable CDP-alcohol phosphatidy]ltransferase"	2 (bases 1 to 11158)		10984043				
/protein_id="AAG05929.1"	2 (bases 1 to 11158)		10984043				
/db_xref="GI:9948599"	2 (bases 1 to 11158)		10984043				
/translation="MISVYOLKPRFONILRPVQRLTORGITANQVTLACLSILVWGALVAFGASHLMLFALVPLMLFMALNAALMGREGKSDGAYINELIDYADSGSLYLPFALLPGVNPPLVYVLLAVISRYAAGNLGTVGASKRYDSPPKSRALCFVYIAGVAGSLPLPLAWIDMLLIALLLITGNRQGLAIVGASRSGDGA"	2 (bases 1 to 11158)		10984043				
complement(846..4511)	2 (bases 1 to 11158)		10984043				
/gene="PA2542"	2 (bases 1 to 11158)		10984043				
complement(846..4511)	2 (bases 1 to 11158)		10984043				
/gene="PA2542"	2 (bases 1 to 11158)		10984043				
/codon_start=1	2 (bases 1 to 11158)		10984043				
/transl_table=11	2 (bases 1 to 11158)		10984043				
/product="conserved hypothetical protein"	2 (bases 1 to 11158)		10984043				
/protein_id="AAG05930.1"	2 (bases 1 to 11158)		10984043				
/db_xref="GI:9948600"	2 (bases 1 to 11158)		10984043				
/translation="MKAKKITLAIYGLLALLLGLAALLCTGSGWILGRPGIDYSGFGIRGAGWAOQLRSLWADQDTLVVERPELRMSFACLAGRLCLOVAERVAIDLPPSSDADSGEISLPTLKLPLSIELGVAAGVSLDNDNOOVAELKLAAMHADGLHIDSVTRGFGDLADLQGLRQPSGDMPLOAEATIGLPAPEGKPMKLAOVGGELOQALTLKGRSSYDGLSGLEQLOLAELNPARLLIRADGFGAGGLPDTLLIATIDASGDKSGYLIRGASLPGEGGVAALSLDGVDAOGADIRQLRADGGEORLELVGAPNRESIAADATLMDKDFPMELIYLAEPPTVTLTFRAEVITQOQRTIAGNSAASGPDAGFTLASPVSGDVLQNLPSLIQIRAGOGAEGVNTLRFNDGVAMDRLQSLSNLPYVAEELRGLSLAPLSGSGRDERLALGVLDVKGRLRQAPLFOARLEGEGORWTLLNLRIDGNKVGSGALDQRLGRIDLINRIGOLMPOLFGDLRDLIAGSLAPORNLNKGQOITAVAEIRAGLIDQANONGTILVAASLREGMDIIFXANRAGDQROOLEDTLOGALLRGLAFDQQLDERGNRGLRAGSELSSGGDKLQOPPARLELADGRVNEGHLCSVGASLSCGGEORLLEPERLTYLDFDLSTKQWPEDPFAQGLNCDLQVLDVPAKGPVAVSLDAGSGTLRLRDQGVNDGYSLRLESLRANRVDSTLDGPKGLGELRVQADIDRPAKPLSGSESLAGLDLSLAPFVPEVETLQGRLNKSGRLSGGLAPRYDGRVILSGNIGGEGLEPTSPENSLDARLAGESELGGGSGRAGGOSLGSNLSGSLTAVNLNIGNRLPYVDVNPYQOLEYEPOLKIGMAQELALSGKYVLPBGDITITLPPSTVKSDDAVIVSGEQPQORPLKMKMDIDVEGQOKLAFSGGUSAEIAGLHHIDNIDTRGILNINSGKRYRAYGQRLTIRARLFFAGITDQPLVIDEAIIRVVDVAVGLRTGSAEPTSAFVSFSPAMSOEQALSTLYLGRMSTGCEDSNMIGEAALAGLAVGSAPLTGEIKR	2 (bases 1 to 11158)		10984043				



```
Db 6290 GCAATCTGACGCGCGCACCCGCTGCTAGCATCCGCCCATGAGACTGACTCAAGG 6231
OY 186 sPhetYrCInserSerInsergly-----GIntrSerAlaIleGlySerse 202
Db 6230 ATTGCTCCGGTCCCGCTGCGGCGCTGCTGCTGAGTACGGCGCTTCGCGGAGAG 6171
OY 202 rHis-----GInlyThrGluProTyraAlaAsnIleLyAlaIleGluGluAsp11 219
Db 6170 CCACCTTAGCGCTGCGGCGTACCCCGCAATGCGCCCTGAAGGCCAATATCGAGGCTTA 6111
OY 219 eThrGInGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGInThrAlaLe 239
Db 6110 TGTGGTAGCCTCGCGCAGCGGACGAAGCGCTTGCAAGCTTTCGCCCAATGCGGA 6051
OY 239 uVal-----AlaAlaArgAlaValaGlyTyrrTyraSpleAspSerIle11 255
Db 6050 AGCCGAGCGGGAAGAGCGCGCCAGCGCTCGCTACTTCCAG----- 6008
OY 255 eArgAsnSerIleGlyGluValaAspValaIleIleHisAsp----- 268
Db 6007 -----GCGCAGATCGAAGCAGAGGTGAAGGAGCGCAAGCCCGCAAGCTCAC 5961
OY 269 -----LeuGlyGluProValTyrrIleAspTyrrArgAlaValaGluValaArg61 284
Db 5960 GCTCAAGGTAGTGGCCGCGGAGCGCTGCGCTCGCCAGGTGAATATACAGTGTCTCG 5901
OY 284 yGIuGlyAlaAspAspLyAlaPheThrThrValaAlaAspGluValaProLeuLeuIleG1 304
Db 5900 CGAGCGCCCGCAGCTCGAGAGACTTCGCTTCGCTCGCGGCAAG---CAGCTCAAAACCGGG 5844
OY 304 yAspValaPheHisIleGlyTyrrGluThrLyAsnLeuLeuIleGluAsnAlaSerAl 324
Db 5843 CGCGAAGCTGAACCACTGCTCTACGAGAGATGCAAGCGCGCTGATCCAGAACCGAGCTTC 5784
OY 324 aGluHisGlyThrPheAspGlyArgTrpLeuAspArgSerAlaAspValaIleLeuProAs 344
Db 5783 GCGCTACGCTTCTTCAGAGCGCGCTTCAGACCAAGCGCTGACATCGATCCGCGCGC 5724
OY 344 pAsnThrAlaAspValaSerLeuIleTyrrAspThrGlyThrGInTyrrArgPheAspGluVa 364
Db 5723 CGCGATCGCGCAGATCGACCTGGTCTACGACAGCGCGCAGCGCTACACCTTCGCGAAGGT 5664
OY 364 lValaPhePheThrIleAspProLySThrAsnGluLeuThrThraspProAspLyLeuPr 384
Db 5663 GAGCTTC-----GACGCGGACTCGATC-- 5642
OY 384 oValLyAsrGluLeuLeuGluGInLeuLeuThrValaMetGlyGluAlaTyrrAsnLe 404
Db 5641 -ATCGACGACAGAACTGTGCGCCCGCATGCTGCGCTTCAGAGCGCGCAACCTATGACTC 5583
OY 404 uGlnAlaValaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrrPheAsnMetValas 424
Db 5582 CGAACTGATCCGCAACTCAACACAGAACTCGAATCCAGCGGCTACTTC----- 5534
OY 424 nThrGluIleValaPheProGluArgGluGInIleGlnAsnAspGlnValaSerPheGluG1 444
Db 5534 ----- 5534
OY 444 nSerSerSerArgTrpThrGluProAlaGlnValaAspLeuSerThrLeuGluProVal11 464
Db 5534 ----- 5534
OY 464 eGluThrValaGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAl 484
Db 5533 -----GAGGCGGTGCGGCTCGACCGCGCGCAGCCGCGGAGCGGAGCG 5493
OY 484 aSerAsnLeuIleGlnAspLyLeuAsnLeuValaAlaAlaLyAlaArgHisLeuTyrrAs 504
Db 5492 C----- 5492
OY 504 pMetProAspAspArgValaLeuAlaIleAsnHisAspAspLyValaAsnArgSerIleLe 524
Db 5491 -----GACGCGCGCGCGCAGCGGATC-- 5471
OY 524 uGlyArgIleSerAspAlaValaSerAlaValaAlaArgAlaIleLeuProAspGluSerG1 544
Db 5471 ----- 5471
OY 544 uAsnGluValaIleAspLeuProGluArgThrAlaLeuAlaAsnArgLySThrProAlaAs 564
Db 5471 ----- 5471
OY 564 pValTyrrGInSerLySTyValaProLeuTyrrValaPheValaAlaSerAspLyProArgAs 584
Db 5470 -----CCGTAAGCGGTGCGCTGCGGAGCGGCGCAAGCCCGGAC 5433
OY 584 pGlyGlnIleGlyLeuGlyTyrrGlySerAspThrGlyThyrArgLeuValaThryrPheG1 604
Db 5432 CATGGCGCTCGCGCTTCGTTCTCGACCGAGCTGGGGCGCGCGGAGTTCAACTGAGAC 5373
OY 604 uHisAsnLeuIleAsnArgAspGlyTyrrGlnAlaGlyAlaGluLeuArgLeuSerGluAs 624
Db 5372 CCGCCATTGGGTGAATGCCGAAGACATAGCGCTGGATTGCGAGGAATCTCCGACAC 5313
OY 624 pLyLyGlyValaLyLeuTyrrAlaThrLyProLeuSerHisProLeuAsnAspGlnLe 644
Db 5312 ACGGACAGACCTCGCGCTGCTGACAGAAATTCCTCGATCCGCGCTCACCGACAAAGCT 5253
OY 644 uArgAlaThrLeuGlyTyrrGInGluValaPheGlyHisSerThrAsnGlyPheAspLe 664
Db 5252 GCGCTTACGACGCGGTACACGTTGAA-----GACCTGGTCGACAC 5211
OY 664 uSerThrArgThrLeu-----GluHisGluIleSerArgSerIleIleGlnAs 680
Db 5210 CGAAGCAGACGCTGACCCCTTGCGCGGAATGCGACAGCAAGCGCGCC----- 5162
OY 680 nGlyGlyTyrrAsnArgThrTyrrSerLeuArgTyrrArgLeuAspLyLeuThyrGlnAl 700
Db 5161 -GATGGCTGGCAGCGCGGTGCTTCGCTGAACTGGAATGCGGAGAGTACAAAGCTG-- 5108
OY 700 aProProGluThrTrpGlnAspLeuProValaAspPheValaAsnGlyLysProSer----- 718
Db 5107 -----GCGCAGCATCGGGCT 5091
OY 719 -GInGluAlaLeuLeuAlaGlyValaAlaValaHisLySThrValaAlaAspAsnLeuValas 738
Db 5090 GAGCAGCTTCCTGATGGCGGAGATCGGCTACTCGCTCGTGAGACCAAGCAAGAGTCTGA 5031
OY 738 nProMetArgGlyTyrrArgGlnArgTyrrSerLeuGluValaGlySerSerGlyLeuValse 758
Db 5030 CCCAGCGCATGGTACCGCCCTGCGCACTTCAAGCTCAAGGGGCGCAAGAAAGGCTCTGTCG 4971
OY 758 rAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValaTyrrSerPheGlyAspAs 778
Db 4970 CGACGCGGACGCTCTCCATGTCGAGCGCATGCGCAAGGCTGACCACTCGCGCGCG-- 4913
OY 778 nAlaTyrrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrrIleTr 798
Db 4912 -----GGCCATCGCTGCTCGCGCGCTCGGAGGTAGGCGGAATCGC 4872
OY 798 pSerAspAsnPheAsnHisValaProTyrrArgLeuArgPhePheAlaGlyGluAspGlnse 818
Db 4871 CACCAACGACATCAAGATCGATCCCGCTCGCTGCTGCTTCCTGCGCGCGCACGAGAG 4812
OY 818 rIleArgGlyTyrrAlaHisAspSerLeuSerProIleSerAspLyGlyTyrrLeuThrG1 838
Db 4811 CGTGGCTGCTAAGATCAACGCGAGCTCTCGCGAGATATCCGATGGCGCAGAGATCGG 4752
OY 838 yGlyGlnValaLeuAlaValaGlyThrAlaGlyTyrrAsnTyrrGluPheMetLysAspLeuAr 858
Db 4751 CGGCGCTTACATGATCCCGCGCAGCGCTGAGTATCAATATCCGCTGCGCGCGCTGCGG 4692
OY 858 gLeuAlaValaPheGlyAspIleGlyAsnAlaTyrrAspLyGlyPheThrAsnAspThrLy 878
Db 4691 CTTGGACCTTCTGTCGACGAGGCGAAGCGCTTCAACTCGCTGACCTTCGCTGATCA 4632
```



Db 937 GTATTGCTCAAGCGCTGACACACTCGGCGACAC-----CAC 975  
Oy 787 GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPhe---AsnHisVal 805  
Db 976 CGCTTTCGCGACCGCTACAGTTTGTCGCACTGCACACCAATGCTTCAACACACACTT 1035  
Oy 806 ProTyrArgLeuArgPhePheAlaGlyIleAspGlnIleTyrGlyTyrAlaHisAsp 825  
Db 1036 CGCCCGCTCCCTCGCTTTCGCGGTCGCCACCAAGTGTGGCGGTTAGCACTACAG 1095  
Oy 826 SerLeuSerProIleSerAspLysGlyTyrIleThrGlyGlyIleAlaValAlaGly 845  
Db 1096 ACCCTCTCGCCGACAGACAGCGACCGTATCGCGCGCGCTACTTGGTGGCAGCG 1155  
Oy 846 ThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIle 865  
Db 1156 AGTCGCGTATCGACAGATTCCTGCTGACCAAAATGGCGGCTCGCGACCTTCGTCGACCA 1215  
Oy 866 GlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyAlaGly 885  
Db 1216 GGCACACTCGTTCACACGCTGACCTGCCAGCCTCAAGACCGTGTGCGCTTGGTGTG 1275  
Oy 886 ArgTyrPheAspProValGlyIleValArgValAspValAlaThrGlyValLysGlu 905  
Db 1276 CGCTGGTATCGCAGTTCGCGCGCTGCGCTGACCTGCCACAGCGCTGATGACGAA 1335  
Oy 906 GlyAsnProIleLysLeuHisPhePheIleGly 916  
Db 1336 GCGGCGC---ATTCGCGCTGCACTTTCATGGG 1365

RESULT 6  
AE006218/c  
LOCUS  
DEFINITION  
AE006218 10302 bp DNA linear BCT 08-MAR-2001  
Pasteurella multocida PM70 section 185 of 204 of the complete genome.  
ACCESSION  
AE006218 AE004439  
VERSION  
AE006218.1 GI:12722230  
KEYWORDS  
SOURCE  
ORGANISM  
Pasteurella multocida.  
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
Pasteurella.  
1 (bases 1 to 10302)  
May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whitam,T.S. and Kapur,V.  
Complete genomic sequence of Pasteurella multocida, PM70  
Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)  
11248100  
2 (bases 1 to 10302)  
Zhang,Q. and Kapur,V.  
Direct Submission  
Submitted (24-OCT-2000) Department of Veterinary Pathobiology,  
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN  
55108, USA

FEATURES  
source  
Location/Qualifiers  
1..10302  
/organism="Pasteurella multocida"  
/strain="PM70"  
/db\_xref="taxon:747"  
complement(70..1830)  
/gene="PM1809"  
complement(70..1830)  
/gene="PM1809"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="AAK03893.1"  
/db\_xref="GI:12722231"  
translation="MKLSITKSLITKVFIFGIVSLTVSLAOLDELQVGIQNTQ  
OLIAHYKGPVATADVLDVLOEATQDEPFLAVTKNIPKSGILNHEHYDSKPKKD  
KILSKGYFDGDFVLSRLEVRSTQAMWRLDYDSGERYRPEVTFENAOJREDYLRN

MINFNKGOPYLINDLSLTNNYTSNFFSVLMQPVLEDEHKITVNDVLLPQPKKNSM  
EYGIQMASDVCPRLOQTATKFWINNRHSERNTLVYSAPOKTEATYKMLKPNKY  
YVEYSAGLSENNKNDTSFASLSAIYIMWHEGWOHSILRVRYSOFLANRKTLL  
LVFPASRYRRTKLDGGLFPTMGDLOKLTIDLGRTWMLSDVDFKMGSSOANRRTLL  
RTVTRLEGLWLRKNTERIIPALRFAGGDRSIRGKXKILAPKRNARLTVGSKRL  
TGSFEYOYVPMWMLATFPADTGAAOQFTKRLRGACMGVMAVVAIKFIDIAIP  
VRDKNRKNIOFYIGLGAEL"  
complement(1980..2609)  
/gene="narP"  
/note="PM1810"  
complement(1980..2609)  
/gene="narP"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="NarP"  
/db\_xref="GI:12722232"  
/translation="MSEKKVLIIDHPLMRGICQLMELAPNEVVDAGNGSEGIS  
GYLKPDEPDALISQRTAQSGLDTLKAALAEIDARILITVSDLSDFITLIDAGD  
ATGLSKQIAGQLFISEETVKYHINLRLKLVHSRVAATVLYFEHNGR"  
complement(2625..5141)  
/gene="cyaa"  
/note="PM1811"  
complement(2625..5141)  
/gene="cyaa"  
/codon\_start=1  
/transl\_table=1  
/product="Cyaa"  
/protein\_id="AAK03895.1"  
/db\_xref="GI:12722233"  
/translation="MNYDLSAQKRYEIDLRLEALSGSSGEPQHVQOLTLTLHI  
NHPNLPGVADAPVGIADPVSPOKQYILITVPSLEANSLLPSTYSINATLGVY  
VMGSIASISQTPKSDLDITWVCHRDLSDEKELQKTHILKMAQFIEINATLGVY  
QKRECFRVAEPLTAENCGSAQYMLIDEPERBAITLAGPLMLHLLEQENENSE  
VERLVTRQOICLDMDVDFGLGOLSANETFGASLMQLYRGIDAQYKSVKILEETYS  
SEYNTVTLIARQKEELLTKLMPSHFDEYLAHQARVYLRKHNELRLGFVRSV  
YLRKTEGMQODPRATNNRLOLQLOLIDEMSDALIELOANMKIKOYKKAHS  
LSEKNLFEVEKSGSKTEKGMVYVNOTPSVAGFVOKRTEYESLNKLVANAYFRIL  
TANDLHITSPNVLTLTRHFVYDLRSLSPVYSSVTNEDLTHACEISLVAAYVIL  
DPRKTIQYKSRIOASDLFSFGPKRESLIGSIDITYRLNMRTRTLNHEGPAILAL  
VAKNMQFEFEERGISLOELHSENELEAGPTALQTEVEEKESALPDSSTYPEID  
ENNNPYKIVORPNTQFYLLOENGKIVYFHSRLMS"  
5284..6222  
/gene="pbp"  
/note="PM1812"  
5284..6222  
/gene="pbp"  
/codon\_start=1  
/transl\_table=1  
/product="Pbp"  
/protein\_id="AAK03896.1"  
/db\_xref="GI:12722234"  
/translation="MVMTRKRSRIATRPSPLAMQANYKACLEKIHAGLYVELVP  
WYTKGVDILTLPKIGKGLFVKELOALIKGADIAVSMKDPHQPAGLMLSTI  
QREDBPDAFVSNYSRTLMELPAGAVGTSIRCOQKOLKRLDPLNHEGWGRTLI  
AKLNGVDATITLAAGLIRIGOOORTSITDYOCCIPAGQCAVIEIROAEIILL  
LAPLADPETTACVLAERAMNRHIOGGQVPIYATVATIEBQULFRALVGEVGSILR  
AEGRSVADAETFLIOVAEKLLDOOGADKILDTLYSN"  
6239..6985  
/gene="hemX"  
/note="PM1813"  
6239..6985  
/gene="hemX"  
/codon\_start=1  
/transl\_table=1  
/product="HemX"  
/protein\_id="AAK03897.1"  
/db\_xref="GI:12722235"  
/translation="MAVIVTRPDVYGAQLTEMLIQAGIFAIHLPLFTTETGREINDLP

```

NRLSQLKAGDYVFAVSKHAVOYAVDTLKNTEGFSMTDLAYFAVGORSAAEFPAQIERA
VLPYRQESSEGLADAMQDSEKHILIRGSGREYSEAOQLGCAKVELECYOR
VPRAYNNAEOISICRAGINRLVITSSEILQYLVDPVPESEHNMKSCOLVTVSQRLA
NLAKOLGHCDDVYVSPRADNVHLIKTLTSLSD"
7026..8429
/gene="PM1814"
/size="8429"
/gene="PM1814"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03898.1"
/db_xref="GI:12722236"
/translation="MAENKITSSESTVPKNEGSTVDKKKTSKGTVPDPEKQAVAOA
PATQPTAKNEVPEKKDAGKMRDEMEKSYKKEPAKQARAESKTAETVYKKGSGGL
ALLALIVAGLACAGYRGLQVDHIOHLSALEOQTNTTPRAKVDLPRTPEQRODIA
OLMRSTRDNEEKIQOLEKELAGKDHSLAQONVNRSAKQOQPDWLTLSDDEL
NNALRKLVDNDVDVGSLSLKTADETLEKMSDPRVMVPTAIINDKQLLAVNADON
VIMORLSQLANIDELVLVDNFDAVNNNEKLDSDMKENAKSATSLNHFIRIT
PNTSDKVLAPNODIYLRENIRLRLOIAIMAVPRODELYKOSLETVASWISYEDT
NSEVAQNFLKTLDELAEOISYVDVPSQLOSLNALDKLDRPSQDVKKIELSVKALVE
ETTESAPSSDASTAEQO"
8442..9707
/gene="hemy"
/size="9707"
/gene="PM1815"
/size="8442"
/gene="hemy"
/codon_start=1
/transl_table=11
/product="Hemy"
/protein_id="AAK03899.1"
/db_xref="GI:12722237"
/translation="MERVLELMVLVLAGLIAGPYLSGKQGYVLIETNSYNIEMSTML
VVEFVYAMAVIGIEVITVRFCLSDGTWHFESRRRAKQKOTLEGIMNMNGSDXS
AKKLGKNAKHSEEPILNFTTAAEAQOCDEPSARRYLEAIELAGSDMLWEAKRT
RILDOQKLPARHSSVDSLVTNKNREYVLAVDIYLSKATLALDKIIDQESAGL
YTAQDFIQDQROVNGLDLDEKNEEGVDGLTWNDQPRRRRLDLVKIALIQRLIDC
NDHEPVEYELTEALKRVEDOSPENRLCTQIARLOEPEDNSKILKMEKRAKNAIEGBH
CCIYRALGYLYVNNDFTOASQAFKHVANKQLEATDVMAAYVEEQVGNALAOV
RESEKETAIVASSLTHSOPEEPVLLSK"
complement(9802..10179)
/gene="rectx"
/size="rectx"
/size="note="PM1816"
complement(9802..10179)
/gene="rectx"
/codon_start=1
/transl_table=11
/product="Rectx"
/protein_id="AAK03900.1"
/db_xref="GI:12722238"
/translation="MOEKGFSEBQIDDTLAFCCQKNMONDRFAENVLTFRAGRGYGF
ARHKOELQOLKGISSLAISBAEOTLARPDQNHNLVLRKKRFPQYAEITDLKQKQWR
YMLSHGFRKQBOAHYGHREYE"
BASE COUNT 3091 a 1951 c 2225 g 3035 t
ORIGIN
Alignment Scores:
Pred. No.: 9,65e-24 Length: 10302
Score: 532.00 Matches: 166
Percent Similarity: 37.768 Conservative: 107
Best Local Similarity: 22.968 Mismatches: 258
Query Match: 11.258 Indels: 192
DB: 1 Gaps: 14
US-09-914-168-2 (1-919) x AE006218 (1-10302)
QY 204 GlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSer 223
DB 1710 CAGAAATACCTCAGCTTAATATGCTGATTAAGCTGATGATTAAGAGAGAA 1651
QY 224 AlameCaspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArg 243
DB 1650 GCCGACGCTTCGAGCGCTCAAGCAATTCGATTAAGCAATTCGATTAAGCAATTCGCG 1591

```

```

QY 244 AlaValGlyTyrTyr-----AspIleAspLeuSerIleIleArgAsnSerIleGlyLys 261
DB 1590 GCATACGCGCTATTATTCAAGAGGAGATGGAAATTTCAATTCGAAGCCAAACCTCTTAA 1531
QY 262 ValAsp---ValIleIleHis---AspLeuGlyGluProValTyrIleAspTyrArgAla 279
DB 1530 AAAGATCAATCAATACCGCCATGTGAGATGAGTCGAACCACTGTGAAATGCGAGATGCGAT 1471
QY 280 ValGluValArgGlyGlyGluAlaAspAspLysAlaPheThrThrValAlaAspGluVal 299
DB 1470 TTTTCTTCGCAAGGTGAAGCCCACTCAAGATCTGAAATTTATCGCCGTGACCAAGAAATTC 1411
QY 300 ProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIle 319
DB 1410 CCAAGAAAA---GGAGAGATGCTCAATATGACCAAACTGTGATTAATTAAGTACATC 1354
QY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAsp 339
DB 1353 CAAAAAATCGCATGTGAGTGCAGGAGTATTCAGCGGATTTTATGATGATGAGTGAACGTTAGAA 1294
QY 340 ValIleLeuProAspAspThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyr 359
DB 1293 GTCAACCTTCACAGCAACCAAGCTTGCGCTTATGATGATGATGATGATGATGATGATGATGAT 1234
QY 360 ArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAsp 379
DB 1233 CGCTTCGCTGAAGTTCCTT----- 1213
QY 380 ProAspLysLeuProValLysArgGlyLeuLeuGlnLeuGlnLeuThrValAsnMetGly 399
DB 1212 ---AANAATGCGCAAAATGCGGAGAGATTAATTCGCTTACATGATTAATTAATTAAGT 1156
QY 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr 419
DB 1155 CAGCCCTACCTGATTAATGATCTTTCTACCTTACCAACAATTAACACCTGCAATATGG 1096
QY 420 PheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGln 439
DB 1095 TTTAGCTCTGTTTAATGACACCTGTTTACATGAGAAGAACCAAAATGTCGATGTCAT 1036
QY 440 ValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThr 459
DB 1035 GTA----- 1033
QY 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479
DB 1033 ----- 1033
QY 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAla 499
DB 1032 -----TTGCTCAACACCACC----- 1018
QY 500 ArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
DB 1018 ----- 1018
QY 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539
DB 1018 ----- 1018
QY 540 ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
DB 1018 ----- 1018
QY 560 LysThrProAlaAspValTyrGlnSerLysValProLeuTyrValPheValAlaSer 579
DB 1018 ----- 1018
QY 580 AspLysProArgAspGlyGlnIleGlyLeuGlyTyrPglYSerAspThrGlyThrArgLeu 599
DB 1017 ---AAAAAATACAGTCAAGTGGGATGCTGCGGATCTGATGATGATGATGATGATGATGATGAT 961

```

```

Oy 600 ValhrLysPheGluHHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu 619
Db 960 CAATTAGCTTGGACAAAGCCTTGCAATTAACATCGGTGCACAGTTCCGATACCAATTTA 619
Oy 620 ArgLeuSerGlnAspLysLysGlyLys---LeuTyrAlaThrLysProLeuSerHis 638
Db 900 TATGTTTTCGGCCAAACAGACCTAGAACCCACCTCAAAATGCTTTATTAATAAAC 841
Oy 639 ProLeuAsnAspGlnLeuArgAlaThrLysGlyGlnGlnGlnValPheGlyHisSer 658
Db 840 CCGCTCACTATTATTACGACTCGCGGGTTTAAAGAAATGAA---AATAAATATGAC 784
Oy 659 ThrAsnGlyPheAspLeuSer-----ThrArgThrLeuGlnHisGlnIleSerArg 675
Db 783 ACAGATCTCTTGGCTCTCTCTATCTGCAATTCGCTATGGAATCATGAA----- 733
Oy 676 SerIleIleGlnAsnGlyLysTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLys 695
Db 732 -----GCAGGTGGCAACATCTTTAGCGTTACGTTGCGTATGAC--- 691
Oy 696 LeuLysThrGlnAlaProProGluThrTyrGlnAspLeuProValAspPheValAsnGly 715
Db 690 ----- 715
Oy 716 LysProSerGlnGlnAlaLeuLeuAla-----GlyValAlaValHisLysThrValAla 733
Db 675 AATGTAAAGACAAACATTCCTGTGTGCCAAGCTGCTGTGACAGCAAGCCGTTTA 616
Oy 734 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753
Db 615 CAAGCGGGCTATTCGCCGCTGGGGGATACACAAACACCATCATGATTAGCGTGT 556
Oy 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyr 773
Db 555 ACATGCTGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
Oy 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyLysIleGln 793
Db 495 ACTTATTTACAGCAC-----CATCGTATCGTTACCCGACATGAA 457
Oy 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
Db 456 CTCGGCTGCTTACATACCAAAATATCGAACATTCACCCGCTGCGTTTTCG 397
Oy 814 GlyLysProGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833
Db 396 GGGGAGATCGCATATGCTGCTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 833
Oy 834 GlyTyrLeuThrGlyLysGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPhe 853
Db 336 GGCAGAAATGATTGCTGCTGCTGCTTATTAACCGGGTCAATTTGAATATCATGATCAAGTG 277
Oy 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe 873
Db 276 TATCCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 217
Oy 874 ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrPheLysProValGlyGln 893
Db 216 ACAAAAGCAATTCACCTATGCGGAGCAATGGGCTACGTTGGCCCTACCCCTAGAGGCA 157
Oy 894 ValArgValAspValAlaThrGlyValLysGlyGlnGlyLysProIleLysLeuHisPhe 913
Db 156 ATTAAGTTTATATTCGCCACCGACGATGAGATTAAGACACAGTAATAATATTCATTTT 97
Oy 914 PheIleGly 916
Db 96 TATATTGCA 88

```

```

RESULT 7
LOCUS AE013668 10831 bp DNA linear BCT 26-JUL-2002
DEFINITION Versinia pestis KIM section 68 of 415 of the complete genome.
ACCESSION AE013668 AE009952

```

```

VERSION AE013668.1 GI:21957368
KEYWORDS Versinia pestis KIM
SOURCE Versinia pestis KIM
ORGANISM Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
REFERENCE 1 (bases 1 to 10831)
AUTHORS Deng W., Burland, V., Plunkett, G., III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Straley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.
J. Bacteriol. 184 (16), 4601-4611 (2002)
JOURNAL 12142430
TITLE Genome Sequence of Versinia
REFERENCE 2 (bases 1 to 10831)
AUTHORS Deng W., Burland, V., Plunkett, G., III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Straley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.
Direct Submission
Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA
FEATURES
source
1..10831
/organism="Versinia pestis KIM"
/strain="KIM"
/db_xref="taxon:187410"
247..990
/gene="cysQ"
/note="y0654"
247..990
/gene="cysQ"
/function="phenotype: central intermediary metabolism:
sulfur metabolism"
/note="affects pool of
3'-phosphoadenosine-5'-phosphosulfate in pathway of
sulfite synthesis; residues 1 to 247 of 247 are 78.45 pct
identical to residues 1 to 246 of 246 from E. coli K12 ;
residues 2 to 247 of 247 are 79.26 pct identical to
(AE008906) affects pool of
3'-phosphoadenosine-5'-phosphosulfate in pathway of
sulfite synthesis [Salmonella typhimurium LT2]"
/codon_start=1
/transl_table=11
/product="sulfite synthesis pathway protein"
/protein_id="AA084242.1"
/db_xref="GI:21957369"
/translation="MMLEQIGLAREACVATIMAYVQGRPLDIAHKKDDSPYATADLA
IAHIDGEPVLGVVYPTVEVNTSAANQAWKECGRRMOIOVDAOPPLVYVSRSH
ACVHMCKQTSYTYPRESFLNPGFRVLSF"
/complement(993..1631)
/gene="y0655"
/complement(993..1631)
/note="residues 26 to 211 of 212 are 59.13 pct identical
to residues 1 to 184 of 184 from E. coli K12 ; B4216;
residues 26 to 212 of 212 are 60.42 pct identical to
(A1627283) putative exported protein [Salmonella enterica
subsp. enterica serovar Typhimurium]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA084243.1"
/db_xref="GI:21957370"
/translation="MKNLPTRGKCPSTLSRISLQGHMKNLSLILSLTPLVLS
AHNIDONRVPVGYRDKGELDYDKTQNFYKKNMSNLSLPGKTRVYQHIAGRTSAKE
MAAPLDAIKQANLPRERYQTTIVTDDAIIITAMVVRNSISSKREFPWSQFIIDS

```



gene  
CDS  
1892..2104  
/gene="y0656"  
/note="residues 1 to 66 of 70 are 77.27 pct identical to residues 14 to 79 of 81 from E. coli K12 : B4217; residues 1 to 66 of 70 are 75.75 pct identical to residues 23 to 88 of 90 from Genpept : >gb|AA123326.1| (AE008906) putative cytoplasmic protein [Salmonella typhimurium LT2]"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAM84244.1"  
/db\_xref="GI:21957371"  
/translation="MRIFORYNPAAKVVAMVYVTLRGRILYIKDGAFFEDNKLIVPRVYKDRKFEVMAEYKNROYLRLQADTLTT"  
/gene="y0657"  
complement(2212..3543)  
/gene="y0657"  
complement(2212..3543)  
/function="putative transport"  
/note="residues 1 to 437 of 443 are 82.15 pct identical to residues 1 to 437 of 447 from E. coli K12 : B4218; residues 1 to 443 of 443 are 81.39 pct identical to residues 1 to 446 of 447 from Genpept : >gb|AA123227.1| (AE008906) putative hemolysin-related protein [Salmonella typhimurium LT2]"  
/codon\_start=1  
/transl\_table=11  
/product="putative membrane protein"  
/protein\_id="AAM84245.1"  
/db\_xref="GI:21957372"  
/translation="MLNSVILILFLIPLAVSAFFSLSEISLASRKIKILKLADEEDITNALRLVLIQTPGMFFVQVIGLNVAALIGGIYDGAASFSEFLPERMAPELAAQSAFVCSFLVLSLTFADLPKRIKGMISPEAVAVRIYNPFRCTIMIFPLVPEFNSMANLIPKRLIMVNRNDITSDSDIYAAVAGAGLAGVLVROEHLIEFPLSEFVPSMSRBNVYIPLRESSEDIKIKISTHRSKRLVCDGHIIDVYGVDSKDLNRYLQMSLYLSSGYIISALVLPDTLLISEALSEFKTAGEDPAVILNEALVGIITLDNVMTTLNGDLVGGQGEQSLVARDENSWLIEGGTPIEDVNRVLHIDFPGSNYETIIGGMVMTLRKIKPRDLPKAYAGAKFEVVDIDYSIKYDILVTRLSDPAIILPKAPHESSDA"  
complement(3821..4459)  
/gene="msra"  
/note="y0658"  
complement(3821..4459)  
/gene="msra"  
complement(3821..4459)  
/function="enzyme; proteins - translation and modification"  
/note="residues 1 to 210 of 212 are 73.33 pct identical to residues 1 to 210 of 212 from E. coli K12 : B4219; residues 1 to 211 of 212 are 72.51 pct identical to residues 1 to 211 of 212 from Genpept : >gb|AA123228.1| (AE008906) peptid methionine sulfoxide reductase [Salmonella typhimurium LT2]"  
/codon\_start=1  
/transl\_table=11  
/product="peptid methionine sulfoxide reductase"  
/protein\_id="AAM84246.1"  
/db\_xref="GI:21957373"  
/translation="MQNDNTAVIDANALFGRLTSLPVSPLAAVGHSMYTIPEGNDLAFFMGFWGAEERLEWQPGVYSTAAQYSGGHTPNPYHEVCSGRGHAENVAVDDPAVYIQQOLQITFENHIDPAQMGQGVGYSAIYVILPEQEDAHKSRFEQOAA MEKADQDRTITSEITVALPFFYAAEDDHOQYLHKRPHGVCIGGICVLPKPV"  
4734..6470  
/gene="y0659"  
4734..6470  
/gene="y0659"  
/note="residues 7 to 578 of 578 are 71.20 pct identical to residues 7 to 577 from E. coli K12 : B4220; residues 7 to 578 of 578 are 71.32 pct identical to residues 7 to 577 of 577 from Genpept : >gb|AA123229.1| (AE008906) putative outer membrane protein [Salmonella typhimurium LT2]"

[illegible]

267 HisAspLeu-----GlyGluProValTyrIleAspTyrArgAlaValGluValArgGly 284  
5019 GCAAAGCTGCTTCCTGCGAGCGGATGATTGATTCGCCGTGCATGATGACTGCAAGGCC 5078  
285 GluGlyAlaAspArgLysAlaPheThrValAlaAspGluValProLeuIleGly 304  
5079 GGTGGCAAAAGCATCTGATTATCAGGCGCTTCGCCGTGATACCCGAAATTGGC 5138  
305 AspValPheHisGlyLysTyrGluThrLysAsnLeuIleGluAsnAlaSerAla 324  
5139 TCGGTCTCAACCATGTTGATTGATTAATTCACCGCTCTCTACTGGTTAGCTTA 5198  
325 GluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspValIleLeuProAsp 344  
5199 CGCGGTGGCTATTTCGACCGCAATATGATCAAAAGCCAGCTCGGTGGCTCAATTGA 5258  
345 AsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluVal 364  
5259 CACGAGCGCTTCGGATGATTGATTTCGACAGTGGCAAGCTTACCGCTTGGTAAGGTT 5318  
365 ValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuPro 384  
5319 ATTTTC-----CAAGGCTCACA 5336  
385 ValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeu 404  
5337 ATCCGAGAGATTATCTGCGAATTTAGTCCGTCATGAGGGGGAATTCATACCTCT 5396  
405 GluAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsn 424  
5397 GATGACATGGCCGAGTTGAACCCCGCTTGGACGCCCAACAGTGTTAACTGGTG--- 5453  
425 ThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln 444  
5454 -----CTTGTTCGCCGAT-----TTCCAGAT 5477  
445 SerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464  
5478 GCCAAACAGAGTAA----- 5492  
465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484  
5493 -----ATTTACCGCTTGAT----- 5507  
485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
5507 ----- 5507  
505 MetProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeu 524  
5507 ----- 5507  
507 ----- 5507  
525 GluArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu 544  
5508 -----CGGTCGTACCCACACGACACTGAG 5531  
545 AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAsp 564  
5532 AATACCGTTC----- 5540  
565 ValTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLysProAlaAsp 584  
5540 ----- 5540  
585 GluGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGlu 604  
5541 -----GAACGTGGGGCGGCTTATGCAACCGATGTCGGTCCGGCTTGACCCGCACTGGCGCT 5597  
605 HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuValGluSerGluAsp 624  
5598 AACCTTTCGATGAACGTCTTCGGGCACTACCTTAACCAACCACTACCGGCTTGGGACCT 5657

625 LysLysGlyValLysLeu-----TyrAlaThrLysProLeuSerHisProLeuAsnAspGln 643  
5658 GACGACACACTGGATTTCACGATACCGTATTCCTTATTAGAAACCCCGTGCAGCACTAT 5717  
644 LeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAsp 663  
5718 TACCTGATACAGGGCGCTTAGACGT-----ACCGACTTAACAGAC 5759  
664 LeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyTyr 683  
5760 ACTAATTCACATACACAGACCTTGAACGTTTGCGACGTCTCTCCACGCGCTGG 5819  
684 AsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGlu 703  
5820 CAAAGCGCGATTAACTCGCTTGGAGCTCGAT----- 5852  
704 ThrThrLysLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu 723  
5853 -----CACTTTAACCAAGCGAGACTACGAGTACCCAGATGTTG 5891  
724 -----AlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArg 741  
5892 CTGATCCGGGGGAGATATTATTCGTACCCGCAACGTGGTGGGGGATGCCGCTCGG 5951  
742 GlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsn 761  
5952 GGGGATAGCCAGCGCTATTATTGATGTCTGATACGACTTGGGGGCTGGATGTTGAT 6011  
762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGly 781  
6012 TTGGCATCTTTTCAGCGCGCAAGACGTGGATCCGACCTTAGTGAGAA----- 6062  
782 SerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGlyTyrIleTyrSerAspAsn 801  
6063 -----AATCGTTTGTGGCCGTGCGCAATGGGGGAGATGCAAAACCAATAT 6110  
802 PheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArgGly 821  
6111 TTGGACCGTGTCCACCGCTGCTCTTCTGCGCGGGGACCGCACTGATCCGTGGC 6170  
822 TyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnVal 841  
6171 TATTAATTCGTAATATTTCACCGCGTGAATGACGAAATTAACCGGCGCTTCCAG 6230  
842 LeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaVal 861  
6231 TTGGCAACGCGCTCTCTGATATCAATATCACTTACACGGGTAGTGGGGCGCTT 6290  
862 PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAla 881  
6291 TTGCTGATTCGGGCGGAGGCGGTGAATTAATTCAGCAAGAGTGACTTAACACAGGGCA 6350  
882 GluValGlyValArgTyrAlaSerProValGlyGlnValArgValAspValAlaThrGly 901  
6351 GGTCTCGGGTACGTTGGGATGATCACCCTTGGACSSATCAAACTGATATCGGGGACCA 6410  
902 ValLysGluGlnGlyAsnProIleLysLeuHisPhePheIleGly 916  
6411 ATA---GGCGATAAAGACACATAGGCGTCAATTTTACATCGGT 6452

RESULT 8  
AJ414157/c 216050 bp DNA linear BCT 06-JUN-2002  
LOCUS  
DEFINITION  
AJ414157 AL590842  
ACCESSION  
VERSION  
AJ414157.1 GI:15981328  
KEYWORDS  
Yersinia pestis.  
Yersinia pestis.  
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Yersinia.  
REFERENCE  
1 (bases 1 to 216050)  
Parkhill,J., Wren,B.W., Thomson,N.R., Tibball,R.W., Holden,M.T.G.,  
AUTHORS

TITLE	Penrice,M.B., Sebalthia,M., James,K.D., Churcher,C., Mungall,K.L., Baker,S., Basham,D., Bentley,S.D., Brooks,K., Cerdeno-Tarraga,A.M., Chillingworth,T., Cronin,A., Davies,R.M., Davis,P., Dougan,G., Feltham,T., Hamlin,N., Holroyd,S., Jagsels,K., Leather,S., Kariyshev,A.V., Mould,S., Oyston,P.C.F., Quail,M., Rutherford,K., Simmonds,M., Skellon,J., Stevens,K., Whitehead,S. and Barrell,B.G. Genome sequence of <i>Yersinia pestis</i> , the causative agent of plague
JOURNAL	Nature 413 (6855), 523-527 (2001)
MEDLINE	21470413
PUBMED	11586360
REFERENCE	2 (bases 1 to 216050)
AUTHORS	Parkhill,J.
TITLE	Submitted (04-OCT-2001) Submitted on behalf of the <i>Yersinia</i> Sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
JOURNAL	
COMMENT	Details of <i>Y. pestis</i> sequencing at the Sanger Centre are available on the World Wide Web. (URL, <a href="http://www.sanger.ac.uk/Projects/Y_pestis/">http://www.sanger.ac.uk/Projects/Y_pestis/</a> ).
FEATURES	1. 216050 Location/Qualifiers
source	/organism="Yersinia pestis" /strain="CO92" /db_xref="taxon:632" /note="biovar: Orientalis"
gene	86..940 /gene="ypj3402" /note="panc"
CDS	86..940 /gene="ypj3402" /note="Similar to <i>Escherichia coli</i> pantocate--beta-alanine ligase PanC SW:PANC_ECOLI (P31663) (283 aa) fasta scores: E(): 0, 72.0% id in 282 aa, and to <i>Vibrio cholerae</i> pantocate--beta-alanine ligase Vc0591 TR:09KUD1 (EMBL:AE004144) (293 aa) fasta scores: E(): 0, 59.0% id in 283 aa"
misc_feature	/codon_start=1 /transl_table=11 /product="pantocate--beta-alanine ligase" /protein_id="CAC92632.1" /db_xref="GI:15981329" /db_xref="SPTREMBL:Q8ZBK7" /translation="MLITETPLRLQOIRRMROGKRIALVPTMGNIHEGAMTLVDEAKTRADVVTITFVNPLOFERPDLAHPTLOECCKLTRHGADLVAPADADYIPAGLEKOTYVDPALSTILEGASRPGHSGVSTVSLFNLIOVDVACFGEKQVOALIRKMYADMGYDINIVGVPTVRAKDGIALSSRMGYLTLEERQILAPQLSKIMMALAEKMAIGERQIDALLEEAQOLRVGFTPDDELTRDAETLOPLTVDSQAVILMAAWLGKARLIDNQLVDLRH"
gene	1351..1731 /gene="ypj3403" /note="pand"
CDS	1351..1731 /gene="ypj3403" /EC_number="4.1.1.11" /note="Similar to <i>Escherichia coli</i> aspartate 1-decarboxylase precursor Pand SW:PAND_ECOLI (P31664) (126 aa) fasta scores: E(): 0, 78.6% id in 126 aa, and to <i>Pseudomonas aeruginosa</i> aspartate 1-decarboxylase precursor Pa4731 TR:AG08117 (EMBL:AE004887) (126 aa) fasta scores: E(): 1.5e-27, 58.2% id in 122 aa"
misc_feature	/codon_start=1 /transl_table=11 /product="aspartate 1-decarboxylase precursor" /protein_id="CAC92633.1" /db_xref="GI:15981330" /db_xref="SPTREMBL:Q8ZBK6" /translation="MIRTMLOGLKLRVKNVQADLVHESCAIDODFLEAGILEYEAIDYVNDGOREFTYVAIAERGSRIISVNGAARACACVGDKLIIISYQMSYAAARLHHPKVAYFEGENOLQKAKAVPYOVA"
misc_feature	1351..1698 /gene="ypj3403"

---

gene	/note="Pfam match to entry PF02261 Asp-decarbox, Aspartate decarboxylase, score 264.30, E-value 1.6e-75" complement(1880) .3073) /gene="ypj3404" complement(1880) .3073) /gene="ypj3404"
CDS	/note="Similar to <i>Escherichia coli</i> hypothetical protein Yade SW:YADE_ECOLI (P31666) (409 aa) fasta scores: E(): 0, 65.4% id in 382 aa"
misc_feature	/codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC92634.1" /db_xref="GI:15981331" /db_xref="SPTREMBL:Q8ZBK5" /translation="MTMRDSRYSVLGEBHYLPVCEIPEGOLIVTPYIAEYERKFGNKGFDKDDLEIKAKRNDMDLGLKPLPQNQILTMREKAVYLADVREQCAVLAEMLRPIVGRKLDRLSNTWYOVNIGDRGVSSADAEIDNIPILTYHMLKNEBKRLNTSTTSVAFSNOMAYLKOAGYDTISLYQLEGYLNKINLPKAVVLTFFDGLKSVSRAYPIILKENGFRATAFIISRIRKRPQKMPNDSLOFNSIAELSGIDYFQIOSHTPLHRTDNRHHPILLSRSYHNIIIFDPEKSRRALQFPNPHVIELSPGCGMOTAVNAAKNGEHLAVTTMCKVKVPCNDPTLKRLLIIRDTSIPTMAERINORCPAVAKPAPA VIEMD"
gene	complement(3400) .4170) /gene="ypj3405" /note="yadh"
CDS	complement(3400) .4170) /gene="ypj3405" /note="Similar to <i>Escherichia coli</i> hypothetical 28.5 kDa protein in hpt-pand intergenic region Yadh SW:YADH_ECOLI (P36880) (256 aa) fasta scores: E(): 0, 80.1% id in 256 aa, and to <i>Vibrio cholerae</i> putative permease Vc0590 TR:09KUD2 (EMBL:AE004144) (256 aa) fasta scores: E(): 0, 66.0% id in 256 aa"
misc_feature	/codon_start=1 /transl_table=11 /product="putative ABC transporter, Integral membrane protein" /protein_id="CAC92635.1" /db_xref="GI:15981332" /db_xref="SPTREMBL:Q8ZBK4" /translation="MTRLYWIALOSIMWIKETTRFARIQITLVPPVITMSLYFIEGNLIGARIGMGGEVDYMOFTVPLGLIMNAVITNAISNVAASFYGAKEFQRSIEELVAVPTHIVIGYGGVGAVGICVGIIVTIIISLFEPVPLHVPVSMISALITLITFAILFSLGILNAAFAKTFPDISVLPFVLTPLTYLGGVYSLSLDPPOAVSKLPIYVIMISGFRYGF LGIDVSLATYITGVLVFIAFYMAAWLIERGTGLRT"
misc_feature	complement(3409) .4164) /gene="ypj3405"
misc_feature	/note="Pfam match to entry PF01061 ABC2_membrane, ABC-2 type transporter, score -22.30, E-value 5.8e-05" complement(3427) .3492) /gene="ypj3405"
misc_feature	/note="one of 6 probable transmembrane helices predicted for YPj3405 by TMHMM2.0" complement(3526) .3639) /gene="ypj3405"
misc_feature	/note="PS00890 ABC-2 type transport system Integral membrane proteins signature." complement(3592) .3657) /gene="ypj3405"
misc_feature	/note="one of 6 probable transmembrane helices predicted for YPj3405 by TMHMM2.0" complement(3679) .3744) /gene="ypj3405"
misc_feature	/note="one of 6 probable transmembrane helices predicted for YPj3405 by TMHMM2.0" complement(3790) .3855) /gene="ypj3405"
misc_feature	/note="one of 6 probable transmembrane helices predicted for YPj3405 by TMHMM2.0" complement(3919) .3964) /gene="ypj3405"
misc_feature	/note="one of 6 probable transmembrane helices predicted

```

misc_feature      for YP03405 by TMHMM2.0"
                  complement(4045. .4110)
                  /gene="YP03405"
                  /note="one of 6 probable transmembrane helices predicted
                  for YP03405 by TMHMM2.0"
                  complement(4167. .5093)
                  /gene="YP03406"
                  /note="yadg"
                  /gene="yadg"
                  complement(4167. .5093)
                  /gene="YP03406"
                  /note="Similar to Escherichia coli hypothetical ABC
                  transporter ATP-binding protein yadg SW:YABG_ECOLI
                  (P36879) (308 aa) fasta scores: E(): 0, 81.8% id in 308
                  aa, and to Vibrio cholerae ABC transporter ATP-binding
                  protein VC0589 TR:Q9KUD3 (EMBL:AB004144) (305 aa) fasta
                  scores: E(): 0, 71.4% id in 304 aa"
                  /codon.start=1
                  /transl.table=11
                  /product="putative ABC transporter, ATP-binding protein"
                  /protein_id="CAC92636.1"
                  /db_xref="GI:15981333"
                  /db_xref="SPTREMBL:Q82BK3"
                  /translation="MTYALEITQLTKTYMGVQVQALRGIDLHVEAGDFALLGPNAGK
                  STTIGLISLVNKTSGVQVGYQFDIDIDYNAKROLGVQENENPFYVLOIVITQ
                  AGYGYTRDLARAERYLSQDLMSKRDRAIRLSGMRRLMIALHBEPLIL
                  DEPTAGVDLELRSMWGLKELNAQGTIIILTHYLEAMLCRNIGIIONGELVENT
                  TMKQLSLSESTFIDLGAKSPLPKLEGYRLTDTSTLEVDVKREOGLNSLFSQIN
                  VQGVQVSMNRKANRLEELFVTLVNGHEGEKA"
                  complement(4458. .5003)
                  /gene="YP03406"
                  /note="Pfam match to entry PF00005 ABC_tran, ABC
                  transporter, score 198.80, E-value 8.5e-56"
                  complement(4641. .4685)
                  /gene="YP03406"
                  /note="PS00211 ABC transporters family signature."
                  complement(4959. .4982)
                  /gene="YP03406"
                  /note="PS00017 ATP/GTP-binding site motif A (P-loop)."
                  5329. .5991
                  /gene="YP03407"
                  /note="yadF"
                  5329. .5991
                  /gene="YP03407"
                  /gene="YP03407"

CDS
Alignment Scores:
Pred. No.:      1.33e-19      Length:      216050
Score:          494.00      Matches:      167
Percent Similarity: 38.04%      Conservative: 105
Best Local Similarity: 23.36%      Mismatches: 255
Query Match:     10.45%      Indels:      188
DB:              1          Gaps:      17

US-09-914-168-2 (1-919) x AJ414157 (1-216050)
OY 211 AsnIleLysAlaAlaLeuGluAspIle---ThrGlnGluSerAlaMetAspLeuAsnGly 229
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136508 AACCTCAGGCGACCTTCTCCAAATTGGCACTGATGAGATTACCGCAGAT-----GCC 136855

OY 230 SerIle---ProAlaLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyr 248
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136854 CCGTTTCGCTTCACGGGTCGATGAGCATATTCCGACGGCGCTCCCGCATTTGATTATTAC 136795

OY 249 Asp-----IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIle 266
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136794 GATCCCAACATACCTTTGAATTCAAAAACCGCCGACCCGACCGACGTTCCAGTCTTAATT 136735

OY 267 HisAspLeu-----GlyGluProValTyrIleAspTyrAlaGlyAlaValGluValArgGly 284
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136734 GCAAAAGTGGTCTCTGGCGAGCCGATATGATTCGCCGCTGCATATCTGACTCAAGCC 136675

OY 285 GluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGly 304
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136674 GGTGCGAAAGCGATCTGATATATCAGGCACCTCGTTCGCCCTGATACCCGAAATTTGGC 136615

```

```

OY 305 AspValPheHisIleGlyLysTyrGlnThrLysLysAsnLeuIleGluAsnAlaSerAla 324
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136614 TCGGTCCCAACACATGGGATTTGATTAATTTCAACAGTCTCTACTGCGTTAGCGTTA 136555

OY 325 GluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspValIleLeuProAsp 344
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136554 CGCCGTGGCTATTTTCAGCGCCCAATATGATCAAAAGCCAGCGGTGGTCCGCTCAATTA 136495

OY 345 AsnThrAlaAspValSerLeuIleTyrAspPheGlyThrGlnTyrArgPheAspGluVal 364
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136494 CACAGCGGCTTCTGGATATTGATTGATTCACAGTGGCAGACCTTACCGCTTGTAAGGTT 136435

OY 365 ValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuPro 384
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136434 ATTTTTC-----CAAGCGTCAACA 136417

OY 385 ValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeu 404
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136416 ATCCGAGAGGATTTATCTGCAGAAATTAGTCCGTTCCATGAGGGGGAATTCCTAATCTCT 136357

OY 405 GlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsn 424
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136356 GATGAACTGGCCGAGTTGAACCCCGCTTGGCAGCCACCACTGTTAACTCGGTG--- 136300

OY 425 ThrGluLeuValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln 444
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136299 -----GTTGTTCCCGCGAT-----TTCCAGAT 136276

OY 445 SerSerSerArgTyrGlnProAlaGlnValAspGluSerThrLeuGluProValIle 464
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136275 GCCAAGAGAGTAA----- 136261

OY 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136260 -----ATTTACCGCTTGAT----- 136246

OY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136246 ----- 136246

OY 505 MetProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeu 524
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136246 ----- 136246

OY 525 GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu 544
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136245 -----GCGTCCGTTACCCCGCAGCGAG 136222

OY 545 AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAsp 564
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136221 AATACCGTC----- 136213

OY 565 ValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAsp 584
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136213 ----- 136213

OY 585 GlyLysIleGlyLeuGlyTyrPglySerAspThrGlyThrArgLeuValThrLysPheGlu 604
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136212 ---GACAGCGGGCGGTTATGACAAACGATGCGTCCGCTGACCCGACAGTTGGCGT 136156

OY 605 HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlyAsp 624
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136155 AAGCCTTGATGAACTCCTTCGGCGATAGCTTAACCAACCACTACCGCGCTTCGGCACT 136096

OY 625 LysLysGlyValLysLeu---TyrAlaThrLysProLeuSerHisProLeuAsnAspGln 643
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136095 GAGCAGACACTGGATTTCAGTACCGTATTCCTTTATTTAGAAACCCCTCGACAGCTAT 136036

OY 644 LeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAsp 663
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136035 TACCTGATACAAAGGGCGCTTAGACGT-----ACGACTTAAACGAC 135994

```

OY	664	LeuSerThrArgThrLeuGlnIstGluSereArgSerIleIleGlnAsnGlyLysP	683
Db	135993	ACTAATTCAGATACACACGACCTTGAACGTTGCACGCTTTTGGGATCTCTCCACGGCGCTGG	135934
OY	684	AsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGlu	703
Db	135933	CMAAGGGCGATTACCTCGCTGGAGCCCTCGAT	135901
OY	704	ThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuLeu	723
Db	135900	-----CACTTTACCCAGGACAGGTCACGGATACCAACGATGTG	135862
OY	724	-----AlaGlyValAlaValAlaHisLysThrValAlaAlaAspAsnLeuValAsnProMetArg	741
Db	135861	CTGTATCCGGGGGTGAGTATTATATGTAATCGTACCCGCCAACGNGTGGGGCGATGCCGCTGG	135802
OY	742	GlyTyrArgGlnArgTyrSerLeuGlnLysValGlySerSerGlyLeuValSerAspAlaAsn	761
Db	135801	GGGGATACCCACCGCTATTTATTTATGATGTCGTATACGACTTGGGGGCGATGTGTTAT	135742
OY	762	MetaIleAlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyrGly	781
Db	135741	TTTGGCATCTTTCACGGCGAGAACGTGAGATCCGACGACTTAGTGACAAA-----	135691
OY	782	SerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsn	801
Db	135690	-----AACCGTTTGTGGCGCGGTGGCAATGTGGGGTGGATGCAAAACCAATAT	135643
OY	802	PheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysArgGlnSerIleArgGly	821
Db	135642	TTTGACCGGTTCACACCGCTCGCTGTTCTTCCCGGGGGGACCGACGATACCGTGGC	135583
OY	822	TyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyVal	841
Db	135582	TATAAATTCGGAATATTTACCCGGGTATAGTAGCAAAATTTAACCGGGCGCTCCAG	135523
OY	842	LeuAlaValGlyThrAlaGlyLysArgLysPheGlyLysPheMetLysAspLeuArgLeuAlaVal	861
Db	135522	TTGGCACGGGCTCTCTTGAATATATATTAATTAATCTTACCGGTAGATGGTGGGGCGCTT	135463
OY	862	PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAla	881
Db	135462	TTCCGTGCGATTCGGCGGACGGCGGTGAATTAATTCACACAGACGACTTAAAAACACGGGCA	135403
OY	882	GlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGly	901
Db	135402	GGTGTGGGGGATACGTTGGCATCACCCGTGGACCGATCAAACTGATATGCGCGACCA	135343
OY	902	ValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGly	916
Db	135342	ATA---GGCGATACGAGACACATGCGCGCAATTTTACATCGT	135301
RESULT 9			
AE008906			
LOCUS			
DEFINITION	Salmonella typhimurium LT2, section 210 of 220 of the complete genome.		
ACCESSION	AE008906		
VERSION	AE006468		
KEYWORDS	AE008906.1 GI:16422950		
SOURCE			
ORGANISM	Salmonella typhimurium LT2.		
REFERENCE	1 (bases 1 to 24892)		
AUTHORS	McClelland, M., Sanderson, K.E., Spieth, T., Clifton, S.W., Latreille, P., Courtillot, L., Portwolk, S., All, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Stewar, N., Mulvaney, E., Ryan, E., Sun, H., Flores, L., Miller, W., Stenking, T., Nhan, M., Waterston, R. and Wilson, R.K.		
TITLE	Complete genome sequence of Salmonella enterica serovar Typhimurium LT2		
JOURNAL	Nature 413 (6858), 852-856 (2001)		

REFERENCE TITLE	AUTHORS	JOURNAL
2 (bases 1 to 24892)	The Salmomella typhimurium Genome Sequencing Project.	
Submitted	(29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA	
COMMENT	Supported by NIH grant 5U 01 AI43283	
COMMENT	Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs: GLIMMER; <a href="http://www.tigr.org/softlab/glimmer/glimmer.html">http://www.tigr.org/softlab/glimmer/glimmer.html</a> and Genemark; <a href="http://opal.biology.gatech.edu/genemark/">http://opal.biology.gatech.edu/genemark/</a>	
COMMENT	EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; <a href="http://www.genome.ad.jp/kegg/">http://www.genome.ad.jp/kegg/</a> , and Pedro Romero and Peter Karp at EcoCyc; <a href="http://ecocyc.PangeaSystems.com/ecocyc/">http://ecocyc.PangeaSystems.com/ecocyc/</a>	
COMMENT	The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and ReguonDB; <a href="http://kinich.cifn.unam.mx:8650/db/regulondb_intro.frameset">http://kinich.cifn.unam.mx:8650/db/regulondb_intro.frameset</a>	
FEATURES	This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.	
SOURCE	Location/Qualifiers	
gene	/organism="Salmomella typhimurium LT2" /strain="LT2; SGSC 1412; ATCC 700720" /db_xref="taxon:99287" /note="LT2" complement(95..383) /gene="yJfY" /note="STM4389" complement(95..370) /gene="yJfY" /note="similar to E. coli orf, hypothetical protein (AAC77156.1); Blastp hit to AAC77156.1 (91 aa), 71% identity in aa 1 - 91" /codon_start=1 /transl_table=1 /product="putative outer membrane protein" /protein_id="AAL23295.1" /db_xref="GI:16422951" /translation="MKRIMLFLALLPGITATNAVELNNHQARNMDVSLGYITINHNFTSEANLALNEEDVRNMYHILIREPSNGNIHASNIYR" complement(378..383) /gene="yJfY" /note="putative RBS for yJfY; ReguonDB:STMS1H004274" 560..770 /gene="STM4390" 560..565 /gene="STM4390" /note="putative RBS for STM4390; ReguonDB:STMS1H004275" 564..770 /gene="STM4390" /note="hypothetical protein" /codon_start=1 /transl_table=1 /product="putative cytoplasmic protein" /protein_id="AAL23210.1" /db_xref="GI:16422952" /translation="MMOKSALTYISLPOQAYNSGRREYVNISAAKCDGTGYFTIINNCITTLSPAGASILIRHFSADL" 828..1236 /gene="trpSF"	
CDS		
gene		
RBS		
gene		
RBS		
CDS		
gene		

```

RBS      /note="STM4391"
828. .833
/genes="rpsf"
/notes="putative RBS for rpsf; RegulonDB:STWSIH004276"
841. .1236
/genes="rpsf"
/notes="similar to E. coli 30S ribosomal subunit protein S6
(AAC77157.1); Blastp hit to AAC77157.1 (131 aa), 99%
identity in aa 1 - 131"
/codon_start=1
/transl_table=11
/product="30S ribosomal subunit protein S6"
/protein_id="AL23211.1"
/db_xref="GI:16422953"
/translation="MRHVEIVFMVHPDSEQVGMIERYSATITGAEGKIHRLDWMGR
VOLADPINKLHKAHYVLMVNEAPQEVIDEITTFRENDVIRSMVRTKHAVTASPM
VAKDERRRRDDEFANETADDAAGDSEE"
1226. .1557
/genes="rplB"
/notes="STM4392"
1226. .1231
/genes="rplB"
/notes="putative RBS for rplB; RegulonDB:STWSIH004277"
1243. .1557
/genes="rplB"
/notes="similar to E. coli primosomal replication protein N
(AAC77158.1); Blastp hit to AAC77158.1 (104 aa), 99%
identity in aa 1 - 104"
/codon_start=1
/transl_table=11
/product="primosomal replication protein N"
/protein_id="AA23212.1"
/db_xref="GI:16422954"
/translation="MTNRLSLGTVCRAPLRKVPSGIPHCQVLEHRSVQEEAGFHR
QAMCGMPVIVSGHENQAITHSITVGSRTIVOGFISCHKAKNGLSKMWLHAEOIELIDS
GD"
1547. .1789
/genes="rpsr"
/notes="STM4393"
1547. .1552
/genes="rpsr"
/notes="putative RBS for rpsr; RegulonDB:STWSIH004278"
1562. .1789
/genes="rpsr"
/notes="similar to E. coli 30S ribosomal subunit protein
S18 (AAC77159.1); Blastp hit to AAC77159.1 (75 aa), 98%
identity in aa 1 - 75"
/codon_start=1
/transl_table=11
/product="30S ribosomal subunit protein S18"
/protein_id="AA23213.1"
/db_xref="GI:16422955"
/translation="MARVPRRRRCFRFTAEQVQELDKDIATLKNYITESGKIVPSRI
TGTAKYOROLRAIKRARIYLSILPYTDHQQ"
1819. .2280
/genes="rplI"
/notes="STM4394"
1819. .1824
/genes="rplI"
/notes="putative RBS for rplI; RegulonDB:STWSIH004279"
1831. .2280
/genes="rplI"
/notes="similar to E. coli 50S ribosomal subunit protein L9
(AAC77160.1); Blastp hit to AAC77160.1 (149 aa), 95%
identity in aa 1 - 149"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L9"
/protein_id="AA23214.1"
/db_xref="GI:16422956"
/translation="MQVILIDKVANLGSIGDOVNVKAGYARNEFLVPEKRAVPATKKNV
EYFEARRAELEAKLADVLAAANARAERKINALEVTIASKAGDEGKLFSGITRDIADA
VTAAGDVAKSEVRLLPNCVLRITGHEHVFQVHSEVFARVILNVAAE"

```

```

gene      2417. .3354
/genes="yifZ"
/notes="STM4395"
RBS      2417. .2422
/genes="yifZ"
/notes="putative RBS for yifZ; RegulonDB:STWSIH004280"
2428. .3354
/genes="yifZ"
/notes="similar to E. coli orf, hypothetical protein
(AAC74606.1); Blastp hit to AAC74606.1 (266 aa), 25%
identity in aa 20 - 250"
/codon_start=1
/transl_table=11
/product="putative permease"
/protein_id="AA23215.1"
/db_xref="GI:16422957"
/translation="MDTORQASPFARKNVVYCAACCLIMGSAVPAIKSGYDLPQIA
TDDIPSKIVPACGRFLPAGGLLFLPLDRLKRRPGRPRPQFRODTLLGLTQSLQTLF
FYIGLAFITSGVKSIMNATGTFEFSVLAAHFTIQNDRLSYNTKGLIGFAGVNVNVS
NGLDSEFNLPBGESVVAALFLSAATLVKRSQTVDPVMTGYQIGGLVLYIGY
VFEGTLTIGHFSSVAITLVLTLLSVAFALMSITLKYNRGMIAPENFLIPVGAALS
AIFLGENTLEMKYMLVLCVSGIMVWVKVR"
1
/genes="yrfB"
/notes="STM4396"
complement(3404. .4039)
/genes="yrfB"
/notes="similar to E. coli orf, hypothetical protein
(AAC77163.1); Blastp hit to AAC77163.1 (224 aa), 85%
identity in aa 13 - 224"
/codon_start=1
/transl_table=11
/product="putative cell envelope opacity-associated
protein A"
/protein_id="AA23216.1"

```

Alignment Scores:

Score:	2.28e-20	Length:	24892
Percent Similarity:	486.50	Matches:	174
Best Local Similarity:	36.90%	Conservative:	112
Query Match:	22.45%	Mismatches:	270
	10.29%	Indels:	219
DB:	1	Gaps:	19

US-09-914-168-2 (1-919) x AE008906 (1-24892)

```

Qy      172  Leupheasnaspcglvalaaslnys-----ValProAglLeuLysAla 185
      :::::  |||  ||:::  ||:::  ||:::
Db      16338  ATTTTCCAGGGGTGGGATATTCACGAGAAATGCGCACATATCCGTCAG 16397
Qy      186  LysPheTyrglnSerSerglnSerglyGluThrSeraIailegLySerHisGlnLys 205
      ::  |||  :::::  :::::  :::::
Db      16398  TTATGTTGGGTAGCTACTGAGCTGAGCAAGCTATCGCGTAGCC----- 16442
Qy      206  ThrGluProTyrlAlaaniIlelYsAlaAlaLeuCluanspllethrIngluSerAlaMet 225
      :::::  |||  :::::  :::::  :::::
Db      16443  -----GCAAATGTTGCTGGAAGCTCAAGGCTATCCGAGAGCTGGAANAAA 16490
Qy      226  AspleuansglySerIleProAglLeuAurGlnThrAlaLeuValAlaAlaArg----- 243
      :::::  :::::  :::::  :::::  :::::
Db      16491  AACGTTGCGGCACAGCTCTAGCATTCACAGAGCATAGGTGACCGCGGATCGGCGCTTT 16550
Qy      244  -----AlaValGlyTyTyraSp--- 249
      :::::  :::::  :::::  :::::  :::::
Db      16551  CGGCGCCGAGTGTATGATTCGATTCGCGAAGGCTTAACGCGTAGCGCTTACGAACCC 16610
Qy      250  ---IleaspleuserIleIleIleArgAsnSerIleclYluValAspValIleIleHisAsp 268
      |||  :::::  :::::  :::::  :::::
Db      16611  ACTATCAAAATTGCATGCTTCCACCGCGTCGAAGACGCGAGATTAATTCGCCAGG 16670
Qy      269  Leu-----GlyGluProValTyrlIleasptyrArgAlaValaGluValArglyGluGly 286
      ::  :::::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      16671  GTTACGCCCGCGGACGCGGTGATTGGCGGCTACAGAGTCAATTTTACCGCGGCGCG 16730

```

Oy	287	AlAspAspLySaLaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal	306
Db	16731	CGCAGCGAAAGATTAATCTGGCCCTTACTGAAACCGCTCG--GCAATCGGCGGG	16787
Oy	307	PheIshIsGlyLyStyGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHis	326
Db	16788	CTGACACGGGAGCATATGCACAATTTTAAAGTCGTTAAACGAGCGTGCTGCTACGTAA	16847
Oy	327	GlyTyPheAspGlyArgTrpLeuAspArgSerValaIleLeuProAspAsnThr	346
Db	16848	GGCTATTTCGACGGCAATCTCTTAAAGTCAGTTGGGAATTCCTGGGCGCATAG	16907
Oy	347	AlaAspValSerLeuIleTyrsPthrcGlyThrGlnTyArgPheAspGluValAlaPhe	366
Db	16908	CGGTTTGGGATATTGATTATATAGCGGTAGCGGTACGCCCTATCGCTTGGACCTGTACACTTC	16967
Oy	367	PheThrIleAspProLysThrAsnGlnLeuThrIhrAspProAspLysLeuProValLys	386
Db	16968	-----GAAGGTTCGCGAGATTCTG	16985
Oy	387	ArgGluLeuLeuGluGlnLeuLeuThrValaAsnMetGlyGluAlaTyraAsnLeuAla	406
Db	16986	GATGAATATTTCACAAATCTGTTGCCGTTTAAAGAGGGGATGAGTACGAATCGAAGAT	17045
Oy	407	ValArgAlaLeuSerAsnAspSerLeuAlaIthrArgTyPheAsnMetValaSnThrGlu	426
Db	17046	TTGGCGGAACATAACCGCGCATTTCCGCTACCGGATGGTTAAATTCGTGTC-----	17096
Oy	427	IleValIhrPheProGluArgGluGlnIleGlnAsnAspGlnValaSerPheGluGlnSer	446
Db	17097	GTGGTGGCGCGCGGA-----TTTGAGAAATCGCGT	17126
Oy	447	SerSerArgThrGlnProAlaGlnValaAspGlnSerThrLeuGluProValIleGluThr	466
Db	17127	AAAACAAAATCTGCTCGC-----	17144
Oy	467	ValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsn	486
Db	17144	-----	17144
Oy	487	LeuIleGlnAspLysLeuAsnLeuValAlaIalalysAlaArgHisLeuTyraSpmetPro	506
Db	17144	-----	17144
Oy	507	AspAspArgValaLeuAlaIleAsnHisAspAspGlyValaAsnArgSerIleLeuGlyArg	526
Db	17144	-----	17144
Oy	527	IleSerAspAlaValaSerAlaValaAlaArgAlaIleLeuProAspGluSerGluAsnGlu	546
Db	17145	-----TTAAAGCGCGGTATCCGCCCGCAACCAAAACAG	17180
Oy	547	ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyra	566
Db	17181	ATTT-----	17183
Oy	567	GlnSerLysLysValProLeuTyraValaPheValaIaSerAspLysProArgAspGlyGln	586
Db	17184	-----GAA	17186
Oy	587	IleGlyLeuGlyTyrPglySerAspThrGlyThrArgLeuValaThrLysPheGluHisAsn	606
Db	17187	ACCGGGGTGGGTACTCCACGACGTCGGGCGCGGTGTAAGCGTCGTGGAAGAAACCG	17246
Oy	607	LeuIleAsnArgAspGlyTyrcGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys	626
Db	17247	TGCGATGACCTTACGGGTATAGCGGTACCGACACACAGACAGATATTCTGCGCGGAACAG	17306
Oy	627	GlyValLysLeu---TyraIalhrLysProLeuSerHisProLeuAsnAspGlnLeuArg	645
Db	17307	GTGCTGGATTTTAGTTTAAATGCGCGTCTGGAAGAAACCCGCTGGAGCAATCTATCTG	17366

OY	646	AATTTTtLeuGlyTYTGInGlnGluValPheGlyHisSerThrAsnGlyPhaspsleuSer	665
Db	17367	GTAACGGCGGCTTTAAACGTAC-----GATTTCGAAGAATACGAGCAGCAC	17414
OY	666	ThArGrThLeuGlnHisGluIleSerArgserIleileGlnAsnGlyLYTrpSmaArg	685
Db	17415	TGCACCACGCGCTT-----GCCGTGCGCGTTACTGTGGACCTTTCCAGTGGATTGCACGGG	17468
OY	686	ThrTyrserLeuArgTyrrArgLeuaspRlysLeuYSThngInAlaProProGluThrtPrp	705
Db	17469	GCGATTAACTGCGGTGGAGTTTTCAC-----	17495
OY	706	GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu-----	723
Db	17496	-----CACTTTTACCCAGGGGAACGTCACCAATATACACCATGCTGTTCAC	17540
OY	724	AlaGlyValAlaValAlaHisLysThryValAlaAspAsnLeuValAsnProMetcArgLytyr	743
Db	17541	CCGGCGCGTAGTATAGCGCAACCCGCTCGCAGGGGAGCATGATCCAACCTCGGGCGCAT	17600
OY	744	ArgGlnArgTyrrSerLeuGlnIuValGlySerSerGlyLeuValSerAspAlaSametla	763
Db	17601	TGCGACGCGTATTACAGTACAGTACATTCGATAGAGCGCTCGGAGCTGGATTTCTCC	17660
OY	764	IleAlaArgAlaGlyIleSerGlyValTyrrSerPheGlyAspAsnAlaTyrglySerAsn	783
Db	17661	GTGCTCGAGCGCCCAAAACGTCGTGATTCGAGACTTTGTACGATCCG-----	17705
OY	784	ArgAlaHisGlnMetHnGlyGlyIleGlnAlaGlyTyrrIeTrpSerAspAsnPhasn	803
Db	17706	-----CACCGTTTTGTGTCATCGCGCCCAATTTGGCGTGATGANAACCGGTGATTTCAC	17759
OY	804	HisValProTyrrArgLeuAlaGrPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrrAla	823
Db	17760	AAAGTTCGCGCCGCGATTACGTTCTTCGCCGGGGCGAACCAGATTCGCGGCTATAAA	17819
OY	824	HisAspserLeuSerProIleSeraspLysGlyTyrrLeuThnGlyGlyGlnValLeuAla	843
Db	17820	TTCAAATCTATTTCCTCTTAAGATAGCGACGGCAATCTTAAAGGCGCCTCAAAACTGGCA	17879
OY	844	ValGlyThrAlaGluTyrrAsnTyrcIuhPheMetLysAspLeuArgLeuAlaValPheGly	863
Db	17880	ACCGGATCGCTGAGTACCAAGTATACGTACACCGGTAATGTTGGGGGCGAGTGTTC	17939
OY	864	AspIleGlyAsnAlaTyrrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal	883
Db	17940	GATACCGCGGAGCGCGGTGATGATTCGCGCTGCAATTCAAACCGGCGCCGGCTC	17999
OY	884	GlyValArgTyrrAlaSerProValGlyGlnValArgValAsp-----ValAlaThrGly	901
Db	18000	GCGGTGCGGTGGCGCTCGCGGTTGGCGCTGTCAAACTGATTTTGGCCGTACCGGTCCGC	18059
OY	902	ValLysGlnGluGlyAsnProIleLysLeuHisPhePheIleGly	916
Db	18060	GACAAGAAGACGACACGCT-----TTACAGCTTTTATATGCT	18095
RESULT 10			
LOCUS	AL627283	249050 bp	DNA linear BCT 06-JUN-2002
DEFINITION	Salmonella enterica serovar Typhi (Salmonella typhi) strain Ctr18,		
ACCESSION	AL627283	AL513382	complete chromosome; segment 19/20.
VERSION	AL627283.1	GI:16505370	
KEYWORDS			
SOURCE	Salmonella enterica subsp. enterica serovar Typhi.		
ORGANISM	Salmonella enterica subsp. enterica serovar Typhi		
REFERENCE	1 (bases 1 to 249050)		1
AUTHORS	Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D.,		
	Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G.,		
	Sebaiha,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T.,		
	Comerton,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N.,		

TITLE	JOURNAL	MEDLINE PUBMED	REFERENCE AUTHORS	TITLE	JOURNAL
enterica serovar Typh1 CT18	Nature	413 (6858), 848-852 (2001)	Farrar,J., Feltwell,T., Hamlin,N., Haque,A., Hien,T.T., Holroyd,S. Jagels,K., Krogh,A., Larsen,T.S., Leathe,A., Moule,S., O'Gaora,P., Parry,C., Quill,M., Rutherford,K., Simmonds,M., Skelton,J., Stevens,K., Whitehead,S. and Bartell,B.G.		
Complete genome sequence of a multiple drug resistant Salmonella	enterica serovar Typh1 CT18	21534947			
2 (bases 1 to 249050)	Parkhill,J.	11677608			
Direct Submission	Submitted (25-Oct-2001)				
Sequencing team, Sanger Centre, Wellcome Trust Genome Campus,	Hinxton, Cambridge CB10 1SA, UK				
E-mail: parkhill@sanger.ac.uk					
Notes:					
Details of S. typhi sequencing at the Sanger Centre are available					
on the World Wide Web.	(URL, http://www.sanger.ac.uk/projects/S_typh1/).				
Location/Qualifiers					
1. .249050	/organism="Salmonella enterica subsp. enterica serovar				
Typh1"	/strain="CT18"				
/db_xref="taxon:90370"	/complement(78..347)				
/gene="STY4596"	/complement(78..347)				
/gene="STY4596"	/codon_start=1				
/transl_table=1	/product="hypothetical protein"				
/protein_id="CAD06718.1"	/db_xref="GI:16505371"				
/db_xref="SPTREMBL:O821H0"	/translation="MRINVLITSLVAGPALGEAHGVCKSQFYANSAANNELTDNTVY				
FKGESISGTHPSIARGMKIIVOOTQADVTDPDSKYTAQLIQKD"	/complement(479..1753)				
/gene="STY4597"	/note="saMb"				
/complement(479..1753)	/note="Similar to salmonella typhimurium cryptic 60-Mba				
plasmid SaMb protein SW:SAMB_SALTY (P23832) (424 aa) fasta	scores: E(): 0, 89.6% id in 423 aa, and to Shigella				
flexner1 ImpB ImpB TR:Q9Z491 (EMBL:A079316) (423 aa)	fasta scores: E(): 0, 71.7% id in 421 aa				
Paralogue of E. coli umuC (UMUC_ECOLI): Fasta hit to	UMUC_ECOLI (422 aa), 63% identity in 420 aa overlap"				
/codon_start=1	/transl_table=1				
/product="UV protection protein"	/protein_id="CAD06719.1"				
/db_xref="GI:16505372"	/db_xref="SPTREMBL:O821G9"				
/translation="MFALADVNSFYASCEKVFPPDLRNTPVYVLSNNDCGVIAIRSKEA	KOLIGIAMVPFOLKTTEPPVAVPVSANVELTYSMSNRMAHLEELAPRVGYSDIE				
MLVDYNIDSCIDEDEFGROLREHVHSGRRTTGVMGPMTAKSAONASKMROFE	GVALTLNCSKPRKTERKLSTLPDVEITGVSGSRITKKLHTMGITITALDILARANPFITKN				
ESVLYERTYRELNGESCISILEEAPPKQDIVCSRSGEVTVTYEAMROAVCOYAEEAA	EXLRBEOFCRHVAIVFKTSFPAVNEPIYGNASEKLTTPDTODTRDI IAAVVALDRI				
VWKGRRACAGCMNDFTPTGISLNLTLPDEVOPHERSEDLMKVIDSIHNSGMCKVMFA	GGSLPEWMOKRELLSPATYFMAEPLVARLR"				
/complement(695..1741)	/gene="STY4597"				
/note="Pfam match to entry PF00817 IMS, ImpB/mucB/saMb	family, score 545.70, E-value 3.1e-160"				
/complement(join(1753..1830,35386..35730))	/gene="STY4598"				
/note="sama"	/pseudo				
/complement(join(1753..1830,35386..35730))					

```

/gene="STY4598"
/EC_number="3.4.21.-"
/note="Similar to Salmonella typhimurium SamA protein samA
SW:SWAMA_SALTRY (P23831) (140 aa) fasta scores: E(): 0,
82.1% id in 140 aa, and to Salmonella typhimurium ImpA
protein ImpA SW:IMPA_SALTRY (P18641) (145 aa) fasta scores:
E(): 1.5e-31, 64.2% id in 123 aa. This CDS has been
interrupted by the insertion of a bacteriophage"
/codon_start=1
/transl_table=11
/product="UV protection protein (pseudogene)"
/db_xref="PSEUDO:CAD06720.1"
/db_xref="REMRMBL:CAD06720"
1831. .1839
/note="phase insertion point. Note this sequence is
duplicated at position 4507585..4507393"
1840. 35385
/note="bacteriophage"
1973. .2350
/gene="STY4599"
1973. .2350
/gene="STY4599"
/note="no significant database hits"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAD06721.1"
/db_xref="GI:16505373"
/db_xref="SPTRMBL:082168"
/transl_table="MNKPLILLLITFSASADKIPSSIONLTAIVYDTFRHSLEDGL
TIRYMKRLIVDAESMFGICNDYVNNKKPGIKITLTLNNTSDGFEINAGGDEC
RKAGMKQKQRLATVYRTSFKPLQ"
complement(2437. .2655)
/gene="STY4600"
complement(2437. .2655)
/note="STY4600"
/gene="STY4600"
/note="Similar to Escherichia coli prophage P2 Ogr protein
SW:OGRK_ECOLI () (72 aa) fasta scores: E(): 4.8e-15, 55.7%
id in 70 aa and to Bacteriophage 186 late gene control
protein B SW:VPB_BP186 (P08711) (72 aa) fasta scores: E():
2.1e-17, 64.2% id in 67 aa."
/codon_start=1
/transl_table=11
/product="putative positive regulator of late gene
transcription"
/protein_id="CAD06722.1"
/db_xref="GI:16505374"
/db_xref="SPTRMBL:082167"
/transl_table="MNKCPKCGSHATRSSEFOYDSTKERYCOCONINGSTFVTHET
VVRIVIPALVNNAPHPHPTVSGGHNMF"
complement(2723. .3823)
/gene="STY4601"
complement(2723. .3823)
/gene="STY4601"
/note="Similar to Bacteriophage P2 late gene control
protein D SW:VPD_BP2 (P10312) (387 aa) fasta scores: E():
0, 54.3% id in 370 aa and to Bacteriophage 186 late gene
control protein D SW:VPD_BP186 (P21679) (389 aa) fasta
scores: E(): 0, 54.7% id in 371 aa"
/codon_start=1
/transl_table=11
/product="putative regulator of late gene expression"
/protein_id="CAD06723.1"
/db_xref="GI:16505375"
/db_xref="SPTRMBL:082166"
/transl_table="MNKSSDRLGINSRSPASFTIEGKDYTTALDARLMSLTITNRC
FADDDIDDDADGDIYLPBRGAVIQQLGKMGKQPLPKGAFVDEIHSASAPRLT
IRRSADFRITLITRREKSMHQTIVGEVVKETIAARNINKVALGKDITDKVDHMQNT
ESASALDKIAROYGAIVASKGNLFLFIQGGGRASGKPLPVITITRKAGGHIFFL
ADGAAVTGYIASVLHTRPKKKEQTVQVRRHKVATPEPEAKOGDYLVGIDENVLV
NRYVATLGNHVRNARKQMRLQGVASFSLQALAEBRADLYTEMPKVGSGFQPIDDAE
WTITITLVHGVSPNCTTSLTLEFKTDITFIE"

```





Oy 627 GlyValIysLeu---TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg 645  
 Db 154086 GTGCTGGATTTTACTTATAAATGCCGCTGTGAATAAACCCGCTGAGACATATCTATCTG 154145  
 Oy 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665  
 Db 154146 GTACAGCGGGGTTTAAACGTACC-----GATTGGAATATACGGAGCAGCAG 154193  
 Oy 666 ThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyTyrAsnArg 685  
 Db 154194 TCGACCAACGCTT-----GCCGTGTGCGCTTACTGGAGCACTTCCACGCCGTTGGCAGCGG 154247  
 Oy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGlnLuhTrp 705  
 Db 154248 GCGATTAACTGGCGGTTCGAGTTTCGAC----- 154274  
 Oy 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu----- 723  
 Db 154275 -----CACTTTACCCAGGAGAACGTCACCAATACCAACCATGCTGTTTAC 154319  
 Oy 724 AlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgLysTyr 743  
 Db 154320 CCGGGCTGATATATACACCGCAACCGCTGCGCAGGGGAGCATGATGCCAACTGGGGCGAT 154379  
 Oy 744 ArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAlaAsnMetAla 763  
 Db 154380 TCGCAGCGCTACTCGGTAGACTTTCGAAATACGGCTGGCGGCTCCGACGTGATTTCTCC 154439  
 Oy 764 IleAlaIleArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn 783  
 Db 154440 GTACTCGACGGCGCAAAACGCTGTGGATTTGGACTTTGATACGATCC----- 154484  
 Oy 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsn 803  
 Db 154485 -----CACCGTTTGTGATCGGGCCCAATTGGGCTGGATAGAAACCGGTGATTTGCAC 154538  
 Oy 804 HisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAla 823  
 Db 154539 AAAGTTCGCGCGGATTTACGTTCTTCGCGGGGGGACGCGAGTATTCGCGGTATTA 154598  
 Oy 824 HisAspSerLeuSerProIleSerAspLysGlyTyrIleThrGlnGlyGlnValIleVala 843  
 Db 154599 TACAAATCTATTTCGCCCTAAAGATACGACGCGCAATCTTAAAGCGCCCTCAAAACTGGGA 154658  
 Oy 844 ValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863  
 Db 154659 ACCGATCGCTGGAGTACACGATATACGTCACCGCAATGCTGGGGGCGGCTTTGTC 154718  
 Oy 864 AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyAla 883  
 Db 154719 GATAGCGCGCAGCGGTACTGATATTTCCCGCTGAGGATTTCAAAACCGGAGCCGGGCTC 154778  
 Oy 884 GlyValArgTTrpAlaSerProValGlyGlnValArgValAsp-----ValAlaThrGly 901  
 Db 154779 GGGCTGCGCTGGCGCTGCGCGGTGGCGCTGCAAACTGATTTTCCGTACCGGTCCGTC 154838  
 Oy 902 ValLysGlnGluGlyAsnProIleLysLeuHisPheIleGly 916  
 Db 154839 GACCAAGACGAACACGCT-----TTACAGTTTATATCGGT 154874  
 RESULT 11  
 AE004323/C  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS

[illegible]



```

Oy 292 PheThrThrValAlaAspGluValProLeuLeuIleGluAspValPheHisGlyLys 311
    |||
    |||
    |||
Db 13904 TTTCACAGCTTTGATCCGCCGCGAGTGTGAGAGTGCATGCCGCGTAAATCAGCTGTG 13845
Oy 312 TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331
    |||
    |||
    |||
Db 13844 TATGCAACCTTCACAAATCGCGGCGATTCGTAACTGCGCTTGCACAAAGCGCTTAAATGCT 13785
Oy 332 ArgTyrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu 351
    |||
    |||
    |||
Db 13784 GATTTTCAGCGCCAGCGCGCTGGAAGTGAATCCAGAGCTGAATCAAGCTCGGCTGATTGTG 13725
Oy 332 IleTyrAspThrGlyThrGluTyrArgPheAspGluValAlaPheThrThrLeuAspPro 371
    |||
    |||
    |||
Db 13724 CACTTGTATAGCGGATCGCTTACCTGTTT----- 13695
Oy 372 LysThrAsnGluLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu 391
    |||
    |||
    |||
Db 13694 -----GCAGCGACCCAGCGCTGGAAGCGATCAAAATCGATGAAATCGCGCTGATG----- 13647
Oy 392 GluLeuLeuThrValAsnMetGlyGluValTyrAsnLeuGluAlaValArgAlaLeuSer 411
    |||
    |||
    |||
Db 13646 TCGGTGCGCCCTTTCACAAACAGCGGAGCTTATTCGTGCGCAAGTCGCGCACTTTAAC 13587
Oy 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431
    |||
    |||
    |||
Db 13586 CAAATATCTTCCAAATACAGATTCGTCTCTTCGCTGTTTGTTGAG-----CCGCGAC 13536
Oy 432 ArgGluGluIleGluAsnAspGluValSerPheGluGluInsSerSerSerArgThrGlu 451
    |||
    |||
    |||
Db 13535 TTATCGCAGCTTG----- 13524
Oy 452 ProAlaGluValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471
    |||
    |||
    |||
Db 13524 ----- 13524
Oy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGluAspLys 491
    |||
    |||
    |||
Db 13524 ----- 13524
Oy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeu 511
    |||
    |||
    |||
Db 13524 ----- 13524
Oy 512 AlaIleAsnHisAspArgGlyValAsnArgSerIleLeuGlyArgIleGlySerAlaVal 531
    |||
    |||
    |||
Db 13523 -----GATGCGGG----- 13515
Oy 532 SerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuPro 551
    |||
    |||
    |||
Db 13515 ----- 13515
Oy 552 GluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGluInsThrLysVal 571
    |||
    |||
    |||
Db 13514 -----CGTGAGTTG 13506
Oy 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGluIleGlyLeuGlyTyr 591
    |||
    |||
    |||
Db 13505 CCGATACAGCTGACCTTACAGCCGCAAGCAGCACTCACTGGAACCCGCTTACGCTAT 13446
Oy 592 GlySerAspThrGlyThrArgLeuValIleLysPheGluHisAsnLeuLeuAsnArgAsp 611
    |||
    |||
    |||
Db 13445 TTCACGCGCTTGAGCTGCTGCTTCAATTAATGCAAGAAACCTTGGGTGAACGCCAA 13386
Oy 612 GlyTyrGluAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyr 631
    |||
    |||
    |||
Db 13385 GGGCACAGTTTGCACAGCACTTCTCTTCTATACAGAGACACCACTTACGGCGGT 13326
Oy 632 AlaThrLysProLeuSerHisProLeuAsnAspGluLeuArgAlaThrLeuGlyTyrGlu 651
    |||
    |||
    |||
Db 13325 TACAGATCCCACTGGAAGATCGTTAAATGAGTACTATTCGATTCAGTACGCGATGAAG 13266

```

```

Oy 652 GluGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671
    |||
    |||
    |||
Db 13265 -----CATTTGCACAG-----CGTATACCGAGAGCTTGGAATCT 13230
Oy 672 GluIleSer-----ArgSerIleIleGluAsnGlyGlyTyrPheAsnArgThrTyrSerLeu 689
    |||
    |||
    |||
Db 13229 AATGATACACTTGAACCGCATTCGCGAGCTTGATGGCGGCGATTCGCGCTCTTAT 13170
Oy 690 ArgTyrArgLeuAspLysLeuLysThrGluAlaProProGluThrThrGluAspLeuPro 709
    |||
    |||
    |||
Db 13169 CGTTATCTCTTAGAATATACGT-----CAGGGTTTACAA 13134
Oy 710 ValAspPheValAsnGlyLysProSerGluAlaLeuLeuAlaGlyValAlaValHis 729
    |||
    |||
    |||
Db 13133 GATGAT-----AACGCCAATTTTATTTGCGCGGAGACTTATACC 13092
Oy 730 LysThr---ValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGluArgTyrSer 748
    |||
    |||
    |||
Db 13091 CGTACGCGGAGCTGCTCCACACAGTGGCTTACTGACTTGGGCGCATTAACAGAGATTACC 13032
Oy 749 LeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGly 768
    |||
    |||
    |||
Db 13031 TTGGAAATATGCGCATCCAGCATTCCTTTCAGAAACACGCGGATTCGCTTACAAACCGGA 12972
Oy 769 IleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGluMet 788
    |||
    |||
    |||
Db 12971 TCTTCTTGCTTACGAACTTATGCGCGCAAT-----CACCGCGCC 12933
Oy 789 ThrGlyGlyIleGluAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
    |||
    |||
    |||
Db 12932 TTAGTTCGCTGATGATGATGAGCCAACTTGCTGATGAATTTACCAACTCTCGCCCTG 12873
Oy 809 LeuArgPhePheAlaGlyGlyAspGluSerIleArgGlyTyrAlaHisAspSerLeuSer 828
    |||
    |||
    |||
Db 12872 CTCGCTTCTTTCGCGGTGTGATTAACAATTCGCGGATACCGGTACCAACAGATTTCA 12813
Oy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGluValLeuValAlaGlyThrAlaGlu 848
    |||
    |||
    |||
Db 12812 CCGCAGAGTCCAGTGGCGGCTTGACAGGCGCGCATACATCCGCCACGCTCAATCGAA 12753
Oy 849 TyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAla 868
    |||
    |||
    |||
Db 12752 TATCAATACCGTTTAAACCGCAACTGTCGCGCGCATTCATGATGATCGGTGATCGG 12693
Oy 869 TyrAspLysGlyPheThrAsnAspThr-----LysIleGlyValAlaGlyVal 885
    |||
    |||
    |||
Db 12692 TTT-----AACGATATCTCGAATGGAAGAAAGGGGTGCGTACGGGATC 12648
Oy 886 ArgTyrPheAlaSerProValGlyGluValArgValAspValAlaThrGlyVal---LysGlu 904
    |||
    |||
    |||
Db 12647 CCGTGGATTTGCGCGCTTGGCCCTATTCGCGTCAATTCGATGGGGGTGATGCGCGG 12588
Oy 905 GluGlyAsnProIleLysLeuHisPhePheIleGly 916
    |||
    |||
    |||
Db 12587 CCGGCGGATGAGTTCAAATTCACCTTCACTTAAAGG 12552

RESULT 12
AX033468 1737 bp DNA linear PAT 21-SEP-2000
LOCUS AX033468
DEFINITION Sequence 1 from Patent WO0047737.
ACCESSION AX033468
VERSION AX033468.1 GI:10280229
KEYWORDS
SOURCE
ORGANISM
Haemophilus influenzae.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
1 (bases 1 to 1737)
AUTHORS Ruelle,J.L. and Thonnard,J.
TITLE Haemophilus influenzae rd outer membrane sequences used as vaccine
PATENT: WO 0047737-A 1 17-AUG-2000;
RUELLE JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ; THONNARD
JOELLE (BE)

```



Db	1561	GATAGTGATTAAGCTGCCGATTAATTACACAGCAAAAGAGCTGCGTTATGGCACAGCGCTT	1620	gene
Qy	864	GLYValAlaGTrpAlaSerProValIGlyGlnValArgValAspValAlaThrGlyValIys	903	CDS
Db	1621	GGTGTGGCTGGCATCGCCAGTGTGGCATTAATTGATATGGCACACCCATTTCGT	1680	
Qy	904	GLUGluGlyAsnProIleIysLeuHisPheheIleIy	916	
Db	1681	GATTAAGATTAACAGCAAAATATTCAATTTTACATCGCA	1719	
RESULT 13				
LOCUS	U32752/C	11833 bp	DNA linear	BCT 29-MAY-1998
DEFINITION	Haemophilus influenzae Rd section 67 of 163 of the complete genome.			
ACCESSION	U32752	L42023		
VERSION	U32752.1	GI:1573692		
KEYWORDS				
SOURCE	Haemophilus influenzae Rd.			
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.			
REFERENCE				
AUTHORS	1 (bases 1 to 11833) Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A., Merricks,J.M., McKenney,K., Sutton,G.G., Fitzhugh,W., Fields,C.A., Gocayne,J.D., Scott,J.D., Shiley,J., Liu,L.I., Glodek,A., Kellay,J.M., Weidman,J.F., Phillips,C.A., Sprלים,T., Hedblom,E., Cotton,M.D., Utterback,T., Hanna,M.C., Nguyen,D.T., Saudek,D.M., Brandon,R.C., Fine,L.D., Fritchman,J.L., Furumam,J.L., Georgaen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M., Smith,H.O. and Venter,J.C. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd Science 269 (5223), 496-512 (1995)			
JOURNAL				
MEDLINE	95350630			
PUBMED	7542800			
REFERENCE				
AUTHORS	2 (bases 1 to 11833) Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S., Borodovsky,M., Rudd,K.E. and Koonin,E.V. Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli Curr. Biol. 6 (3), 279-291 (1996)			
JOURNAL				
MEDLINE	96398784			
PUBMED	8805245			
REFERENCE				
AUTHORS	3 (bases 1 to 11833) White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D. Direct Submission Submitted (23-JUL-1995) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL				
MEDLINE	4 (bases 1 to 11833) White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D. Direct Submission Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL				
MEDLINE	5 (bases 1 to 11833) White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D., Peterson,J., Hickey,E., Dodson,R. and Gwinn,M. Direct Submission Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL				
MEDLINE	On Sep 30, 1996 this sequence version replaced gi:1221393. Location/Qualifiers 1. 11833 /organism="Haemophilus influenzae Rd" /db_xref="taxon:71421"			
REMARK				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
JOURNAL				



[illegible]

LOCUS	243184 bp	DNA	linear	BCT 07-MAR-2001
DEFINITION	Escherichia coli O157:H7 DNA, complete genome, section 19/20.			
ACCESSION	Ap002368 BA000007			
VERSION	Ap002368.1 GI:13364484			
KEYWORDS	<p>Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RIMD 0509952) DNA.</p> <p>Escherichia coli O157:H7</p> <p>Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.</p>			
ORGANISM	Escherichia coli O157:H7			
REFERENCE	1 (sites)			
AUTHORS	Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T., Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shinagawa, H.			
TITLE	Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak			
JOURNAL	Genes Genet. Syst. 74 (5), 227-239 (1999)			
MEDLINE	20198780			
REFERENCE	2 (sites)			
AUTHORS	Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M., Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and Hayashi, T.			
TITLE	Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an Escherichia coli K-12 strain MG1655			
JOURNAL	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)			
MEDLINE	20557356			
REFERENCE	3 (sites)			
AUTHORS	Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S., Yutsudo, C.H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T., Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shinagawa, H.			
TITLE	Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak			
JOURNAL	Gene 258 (1-2), 127-139 (2000)			
MEDLINE	20564182			
REFERENCE	4 (sites)			
AUTHORS	Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shibb, T., Hattori, M. and Shinagawa, H.			
TITLE	Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12			
JOURNAL	DNA Res. 8 (1), 11-22 (2001)			
MEDLINE	21156231			
REFERENCE	5 (bases 1 to 243184)			
AUTHORS	Ohnishi, M., Kurokawa, K., Makino, K., Yasunaga, T., Shinagawa, H. and Hayashi, T.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail:ken@gen-info.osaka-u.ac.jp, URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365, Fax:81-6-6879-2047)			
MEDLINE	21156231			
REFERENCE	genome project.			
COMMENT	Location/Qualifiers			
FEATURES	<p>1. .243184</p> <p>/organism="Escherichia coli O157:H7"</p> <p>/strain="O157:H7"</p> <p>/sub_strain="RIMD 0509952"</p> <p>/db_xref="taxon:83334"</p> <p>466.. 2115</p> <p>/gene="ECs5008"</p> <p>466.. 2115</p> <p>/gene="ECs5008"</p> <p>/note="similar to pGL-ECOLI g11790457 percent identity 100 in 549 aa (Conserved in E.coli K-12)"</p> <p>/codon_start=-1</p>			
CDs	gene			



/evidence=not\_experimental  
 /transl\_table=11  
 /product="glucosephosphate isomerase"  
 /gene="ECs5009"  
 /protein\_id="BAB38431.1"  
 /db\_xref="GI:13364485"  
 /translation="MKNIIPQTAAMQALQKHFDEMKDVTIADLFARDGDRFSKFA  
 FDDMLVYSKRIRTEETLAKLODLAKCEDLACATSMFSGEKINRPNENRVLHVALR  
 NRSNTEPLVDCGDVEVVALVEKMTFSEAITISGEMKGTGKATIDVNINGTIGSDG  
 GMYATELIRPKKNMLNFVNSVDSTHNAVEYLKKNPETTLVLVASKTGTSTPTMNR  
 AHSARDMLKAAGDRHKVHKHFRALSTHNAKAVEFGIDTANMEFMDVGVSGRYSLSA  
 IGLSVLSIGDNPVELLSGAMKDNHFTTPAEKMLPVLLALIGIWNFFGAETEA  
 ILIPDVAHFRFENFVQOGMESNGKVDNRNGNVVDYOTGPIINGEPTNGOAHAYOLIL  
 HOGKMRVPCDFLAPATITTHNPDLSDHOKLISNFPADTEALAFKRSREVEDEYRQOGD  
 PATLDVYVPEFGFEENRPNRNLIRPTPEPSALIALAYEHRKIFTGQVILNIFFDQ  
 GVLELQOLANRLIPELKDDKEISSHDSSTNGLINRYKAMRG"  
 2628. 2870  
 /gene="ECs5009"  
 2628. 2870  
 /gene="ECs5009"  
 /note="similar to yJBE\_ECOLI g11790458 percent identity  
 100 in 80 aa (conserved in E.coli K-12)"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAB38432.1"  
 /db\_xref="GI:13364486"  
 /translation="MKKVLVYGFATISALATSAMAAPVOYGEAASNAATSVASGSSA  
 TSVSTSVSSAGVALATGSGDGSNTGTTTSTG"  
 2954. 3622  
 /gene="ECs5010"  
 2954. 3622  
 /gene="ECs5010"  
 /note="similar to yJBE\_ECOLI g11790459 percent identity  
 96 in 222 aa (conserved in E.coli K-12)"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAB38433.1"  
 /db\_xref="GI:13364487"  
 /translation="AVIISPLRRVRYKRPAPFIIITLLOCSATTTKGLNSLMSLT  
 PGVOLTDDIDNMPAVISQYMOALNGOLFVLLAFARDGQKWTODATLYTQGRVLY  
 KTLIGDNLIEYNNLAADPLIRPAQITDGTATMKRTYKGTETQOVRATATASVFRWGDG  
 GTVAVGSGEIAIRVLDVEYSTDOARWNRMYWIDSGQIKOSEQYLGADYRPVKTTLLIK  
 AAKQ  
 3619. 4356  
 /gene="ECs5011"  
 3619. 4356  
 /gene="ECs5011"  
 /note="similar to yJBE\_ECOLI g11790460 percent identity  
 97 in 245 aa (conserved in E.coli K-12)"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAB38434.1"  
 /db\_xref="GI:13364488"  
 /translation="MIKOTIVALLSGASSVFAAGTVKVFSGSSSEAKLTUGAEHLI  
 DLVQSPRLANSWMPGAVISEELATTAALRQOQALITRLAEQGDSSADMAALNALRQ  
 OIALKVTGRKKNINDPQIVRAEERANPPLQGNITLTVGCPPESTVTLFGILSRGNDP  
 FTRPDAVASYISGGSLGEGADSVAMVVPDGRTORAPAVAYMNRHVEPMSGITIYVG  
 LADVSWSEIARVLDIOTLRIPO"  
 4356. 6452  
 /gene="ECs5012"  
 4356. 6452  
 /gene="ECs5012"  
 /note="similar to yJHB\_ECOLI g11790461 percent identity  
 99 in 698 aa (conserved in E.coli K-12)"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="hypothetical protein"

```

/translation="MKRHHLSLALGISTACVGETYPAIPGPSODSGEGYLLQPT
/db_xref="GI:13364489"
/protein_id="BAB38435.1"
/gene="ECS5013"
/feature="ECS5013"
/note="similar to yJBA_ECOLI g111790462 percent identity
98 in 136 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/translation="hypothetical protein"
/protein_id="BAB38436.1"
/db_xref="GI:13364490"
/translation="MSTLSRPRVEFISTITLOTVNLGLLGLTLVFLGKPYHLAD
VLFAPDSKVELVEGLVVFYEPIALTYKQSFPHPLRFVYIGITAIYRLII
VDKSPIDVLIYSAILLIVTTLMLCKSRKRR"
complement(7501..8976)
/gene="ECS5014"
complement(7501..8976)
/gene="ECS5014"
/note="similar to xYLE_ECOLI g111790463 percent identity
100 in 491 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/translation="xylene-proton symport"
/product="xylene-proton symport"
/protein_id="BAB38437.1"
/db_xref="GI:13364491"
/translation="MNTOTNSSFITTLVATLGLGLFGDPAVIGSTVESLNTVFA
PQNSSEANSLDGEVCSALIGCTIGGALGYCSNFRGRPSLKAIAVLFISGVSS
AMEPLGTSINPDMLTVYVLAGVPEPVIIRIIGIGLGLAMSLMYPALFPAHJR
GKLVSENPQAIIFGOLLVYCVNFTIARSGASMLNTGMRMSESECIFALFMLTY
TVSPSRMISRCRQBOAEGLIKRIMKNTALTOAVDGIKISLKHGRGTRGLFMFGG
VIVIGWMLSFQOPFVGINVLYIAPVFKTLGASTDIALLOTIIVGINTPLVLAIM
TVDFGRKPLDITIGALMAIGMFLGTAFTYDAGVIALISMLETYVAFMMSGVCW
VLJSEIPFIAIRKALAIIVAAQWLAVFPMSPMDKMSVIAHFHNGFSYIVGC
MGVLAIEPMKRFVPEYTGKTLLELEALMPEBTKKTOOTATL"
complement(9348..10238)
/gene="ECS5015"
complement(9348..10238)
/gene="ECS5015"

gene
CDS
db 211481 CGCGACGCGTCGATGATGTCATCCGGAAGCTCTGAAGACCGACCGCTT
231 ILEPQATGLeuATgGlnThAlaLeuValAlaAlaArgAlaValAlGlyTYTyr-----
248
Alignment Scores:
Pred. No.: 8,77e-16 Length: 243184 *
Score: 434.50 Matches: 164
Percent Similarity: 36.92% Conservative: 100
Best Local Similarity: 22.94% Mismatches: 262
Query Match: 9.19% Indels: 189
Gaps: 18
US-09-914-168-2 (1-919) x AF002568 (1-243184)
OY 211 AsnileysalaalaaleuGluaSPlelThcIngluSerAlaMetaspLeuAsnglySer 230
Db 211421 AACGTCGTGGGACGCTTTCACGATTGAAGTGATGAAGACCGACCGCTT 211480
OY 231 ILEPQATGLeuATgGlnThAlaLeuValAlaAlaArgAlaValAlGlyTYTyr----- 248
Db 211481 CGCGACGCGTCGATGATGTCATCCGGAAGCTCTGAAGACCGACCGCTT 211540

```

Oy 249 -----AspIleAspLeuSerIleIleArgAsnSerIleGlyValValAlaIleIle 266  
 Db 211541 ACCATTGAATTTGATCTCGCTCCACCGCCAAAGAGGGCGAGATTGATGCGCAAA 211600  
 Oy 267 HisAspLeuGlyGluProValIleAspTyrArgAlaValGluArgIleGlyGly 286  
 Db 211601 GTCACGCCAGCGCGTGTAAATTTGGCGCACCGATGTGTTATTTGCGCGGCGCG 211660  
 Oy 287 AlaAspAspLeuAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal 306  
 Db 211661 CGGACCGATTAAGACTATTGTAATTTGCTCGATCTACGCCCG---CCTATTGGCACCGTA 211717  
 Oy 307 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326  
 Db 211718 CTGAACACGCGCATTAATGAAATTTCAAAAGTCTTAACAGCATGCGCTGCGCTAAA 211777  
 Oy 327 GlyTyrPheAspGlyArgTrrPheAspArgSerValAspValIleLeuProAspAsnThr 346  
 Db 211778 GGTATTTCGATAGCGAATTTACCAAGCGCAGCGCGCATGCGCTGCGCTGCTATAA 211837  
 Oy 347 AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhe 366  
 Db 211838 GCGTTCTGCGATATTGATTATTAACACTGCGCAACGCTTACCGCTTGGCATGTGACCTTT 211897  
 Oy 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386  
 Db 211898 -----GAAAGATCACAAATCCCG 211915  
 Oy 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406  
 Db 211916 CATGAATTAACCTCAAAATTTGCTGCGCTTAAAGCGCGCATGAGTACCAATTCGAAAGAT 211975  
 Oy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlu 426  
 Db 211976 CTGGCAGAACTGAACCGCGCATCTCTCTACCGCTGTTAACTCGGCG----- 212026  
 Oy 427 IleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluInSerSer 446  
 Db 212027 GTGGTGGCTCCACAA-----TTTGATTAAGCGCGC 212056  
 Oy 447 SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThr 466  
 Db 212057 GAACGAAAGATTATACCA----- 212074  
 Oy 467 ValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsn 486  
 Db 212075 -----TTTCACG---GCGCTG-----GTTTCGCGC----- 212095  
 Oy 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506  
 Db 212095 ----- 212095  
 Oy 507 AspaAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526  
 Db 212095 ----- 212095  
 Oy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546  
 Db 212096 -----CGAACAGAAACACT 212110  
 Oy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566  
 Db 212111 ATC----- 212113  
 Oy 567 GlnSerLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586  
 Db 212114 -----GAA 212116  
 Oy 587 IleGlyLeuGlyTrrPglySerAspThrGlyThrArgLeuValThrLysPheGluHisAsn 606  
 Db 212117 ACCGGGCTCGTTACTTACGAGACGCGGAGCGCGCTGAAGCGACGCGAAAAAACCG 212176

Oy 607 LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys 626  
 Db 212177 TGGATGAACATCAAGCTGACAGCTGACACCAACAGTACAGTATTTCCGCCGAAACAG 212236  
 Oy 627 GlyValLysLeu---TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg 645  
 Db 212237 ACCCTGACCTTCACTAATAAATGCCCTGCTGAAGATTCACAGTACATATTATTTC 212296  
 Oy 646 AlaThrLeuGlyTyrGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeuSer 665  
 Db 212297 GTCGAGGCGGTTTTAAGCCACT-----GACCTGAACGATACCGAGCTTGAC 212344  
 Oy 666 ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyTrrPasnArg 685  
 Db 212345 TTCACACTACGCTG-----GTGGCTTCTCGCTACTGGGATGCTCCACGCGCTGGCACGCT 212398  
 Oy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrp 705  
 Db 212399 GCCATTAACTGCGCTGAGCTTCGAC----- 212425  
 Oy 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnAlaLeuLeu----- 723  
 Db 212426 -----CACTTACTCAGGTTGAATTAACCAACACGATGCTTTTAT 212470  
 Oy 724 AlaGlyValAlaValHisLysThrValAlaAlaAspAsnLeuValAsnProMetArgGlyTyr 743  
 Db 212471 CCTGGGCTGATGATTAGCCCGCACGCGTCTGCTGGCTGATGCCAACCTGGGCGAC 212530  
 Oy 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763  
 Db 212531 TCCCAACGCTACTCTATCGACTCTCCACACAGCGCTGGGGGTTGAGTGTGCGATTTCG 212590  
 Oy 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTrrGlySerAsn 783  
 Db 212591 GTTTTCAGCGCGAGAACGCTGATGCCACACTGACATGCGC----- 212635  
 Oy 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTrrIleTrrPserAspAsnPheAsn 803  
 Db 212636 -----CATGCTTTTGTACACCGCGCACGCTGGCGGATTAACACCGCTGATTTCGAC 212689  
 Oy 804 HisValProTyrArgLeuArgPhePheAlaGlyIleAspGlnSerIleArgGlyTrrAla 823  
 Db 212690 AAGGTACCGCGGATCTGCGTTTCTTCCGCGGGCGCATGCGATTTCCGCGCTACAAA 212749  
 Oy 824 HisAspSerLeuSerProIleSerAspLysGlyTrrLeuThrGlyGlyGlnValAla 843  
 Db 212750 TACAAATCTATCCGCCGAAATACGCTAACGSGTAGCTGAAGGGCGCTCGAAGTTGATA 212809  
 Oy 844 ValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863  
 Db 212810 ACCGATCGCTGGATTAACAGTACACAGTGAACCGGAAATGTGGGCGGCTGTTGTC 212869  
 Oy 864 AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883  
 Db 212870 GATAGTGGCGAAGCGGTAACCGATTTCCGCGCGACACTTTTAAACCGGTACCGGGCTC 212929  
 Oy 884 GlyValArgTrrPalaSerProValGlyGlnValArgValAspValAlaThrGlyVal--- 902  
 Db 212930 GGGCTACGCTGGGAATCGCGCGTGGGCGCAATCACTCATTTTGGCGTAACGGTGGCG 212989  
 Oy 903 ---LysGluGluGlyLysnProIleLysLeuHisPhePheIleGly 916  
 Db 212990 GATTAAGACGAAACACGG-----TTACAGTTTATCATTCGCT 213025  
 RESULT 15  
 LOCUS AE000493 10819 bp DNA linear BCT 01-DEC-2000  
 DEFINITION Escherichia coli K12 MG1655 section 383 of 400 of the complete genome.  
 ACCESSION AE000493 U00096  
 VERSION AE000493.1 GI:2367360  
 KEYWORDS Escherichia coli K12.

ORGANISM	Escherichia coli K12	CDS	156..656
REFERENCE	Bacteria: Proteobacteria; gamma subdivision: Enterobacteriaceae; Escherichia.		/gene="ytfI"
AUTHORS	1 (bases 1 to 10819)		/function="orf; unknown"
	Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B. and Shao,Y.		/note="Q166"
TITLE	The complete genome sequence of Escherichia coli K-12		/codon_start=1
JOURNAL	Science 277 (5331), 1453-1474 (1997)		/transl_table=11
MDLINE	97426617		/product="orf; hypothetical protein"
REFERENCE	2 (bases 1 to 10819)	gene	/protein_id="AAC77172.1"
AUTHORS	Blattner,F.R.		/db_xref="GI:1790661"
TITLE	Direct Submission		/translation="MLPRLRNHFIGAVELFVKSSYTKHNSNPFNNJHAFKKDWI
JOURNAL	Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.	CDS	SNYSDLTLLREFRCATQIDKSGYOVLSKNETVHAMDKFLISFSLKDGATYMTLR
	Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		GSGEFEELPITINEXNSFMDKRNREPLEQDDCPGIFYKKNLSUTYRKGLKVIETR
REFERENCE	3 (bases 1 to 10819)		RLVTCI"
AUTHORS	Blattner,F.R.		complement(1156..1710)
TITLE	Direct Submission		/gene="ytfJ"
JOURNAL	Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.		/note="b4216"
	Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		complement(1156..1710)
REFERENCE	4 (bases 1 to 10819)		/gene="ytfJ"
AUTHORS	Plunkett,G. III.		/function="orf; unknown"
TITLE	Direct Submission		/note="f184: 100 pct identical amino acid sequence and
JOURNAL	Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA		equal length to YTFJ_ECOLI SW: P39187"
	On Sep 9, 1997 this sequence version replaced gi:1790660.		/codon_start=1
COMMENT	This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 (e-mail: markborov@gatech.edu). Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible ( <a href="http://cgsc.biology.yale.edu">http://cgsc.biology.yale.edu</a> ). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site ( <a href="http://www.genetics.wisc.edu">http://www.genetics.wisc.edu</a> ). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.	promoter	/transl_table=11
		promoter	/product="orf, hypothetical protein"
		promoter	/protein_id="AAC77174.1"
		promoter	/db_xref="GI:1790663"
		gene	/translation="MAPHRNRNQOETMKIFQRYNPLQVAKYKILFRGLYIKDVA
		gene	FEFDGKILIPVKDKLHLSVNSEVNRQVMRLQTEMA"
		CDS	complement(2320..3663)
			/gene="ytfL"
			/note="b4217"
			complement(2320..3663)
			/gene="ytfL"
			/note="b4218"
			complement(2320..3663)
			/gene="ytfL"
			/function="putative transport; Not classified"
			/note="f447: 100 pct identical amino acid sequence and
			equal length to YTFI_ECOLI SW: P39319"
			/codon_start=1
			/transl_table=11
			/product="putative transport protein"
			/protein_id="AAC77175.1"
			/db_xref="GI:1790664"
			/translation="MLNSLIVILCLINSAFSPMSSEISIAARIKKLKLDAGNINA
			ORVLMQENPGMEFTVVOIGLNAVALIGDIVDAFSPAFSLFSRYMSAEISEQLSF
			ILSFSLVGMFTLFADLPKRGIMLAPENVALRIINPMRFCLYCTPLWPFNGLANI
			IFRIFKLPMVRKDDITSDDIYAVAGALAGVLRQOEHLIENVELESRYVPSMTPT
			RENVIWFDLDEDSILKKNVAHPHPSKFLACVDEDIDHITIGYVDSKDLNRLVANSIA
			LNSGVOINNTLIVPDLTLSEALSESFKTAGEDPAVIMNEYALVGCITITLNDVMTLNG
			DLVGGLDEQIYARDBNSMLIDGGRPIIDVMYVDIDEPQSGNVEITIGCFPMMLRK
			IPKRTDSVFAFGYKFEVVDIDNTRIDQLLVTRIDSKATRLSPDKADKESVA"
			complement(3722..3750)
FEATURES	Location/Qualifiers	promoter	
source	1..10819		
	/organism="Escherichia coli K12"		
	/strain="K12"		
	/sub_strain="MG1655"		
	/db_xref="taxon:83333"		
protein_bind	36..56		
	/note="central position to predicted promoter: -73..5"		
	/bound_moiety="IlyV predicted site"		
promoter	86..114		
	/note="factor Sigma70; predicted +1 start at 443250"		
gene	156..656		
	/gene="ytfI"		
	/note="b4215"		

```

promoter
  /note="factor Sigma70; predicted +1 start at 4438844"
  /complement(3785..3813)
  /note="factor Sigma70; predicted +1 start at 4438907"
  3905..3938
  /note="REP (repetitive extragenic palindromic) element;
  contains 1 REP sequence"
  /complement(3986..4624)
  /gene="msrA"
  /note="b4219"
  /complement(3986..4624)
  /gene="msrA"
  /function="enzyme; Proteins - translation and
  modification
  /note="f12; 100 pct identical amino acid sequence and
  equal length to PMSR_ECOLI SW: P27110"
  /codon_start=1
  /transl_table=11
  /product="peptide methionine sulfoxide reductase"
  /protein_id="AAC77176.1"
  /db_xref="GI:1790665"
  /translation="MSLPDKKHLVSPDALPGHNPMPVATLHAVNGHSMTNVPDGM
  IAFAMGCFMGVLEFLMQLPGVYSTAGTGGTTPMTREVCSGDTGAAEARIYVD
  PSYISYELQVEMENHDPAQROGNDHGTQYRSALYLPDPDAAARASLERFQAA
  MIAADDDRHITTEINATPFIYAEDDHQYLHKNPYCGIGIGVCLPPEA"
  4662..4690
  /note="factor Sigma70; predicted +1 start at 4439826"
  /complement(4687..4718)
  /note="factor Sigma70; predicted +1 start at 4439809"
  4751..4762
  /note="central position to predicted promoter:61"
  /bound_molety="AraC predicted site"
  4830..6563
  /gene="yrfM"
  /note="b4220"
  4830..6563
  /gene="yrfM"
  /function="orf; Unknown"
  /note="0577; 100 pct identical amino acid sequence and
  equal length to YrfM_ECOLI SW: P39320"
  /codon_start=1
  /transl_table=11
  /product="orf, hypothetical protein"
  /protein_id="AAC77177.1"
  /db_xref="GI:1790666"
  /translation="MRTIROLCCVSLICSSAVANRFLQYEGLSGLEKVRQALS
  TIESDEVPRFRARVDALREGLKALGTYOPTLEFDLRPPKRGVLIATKVTGV
  PVLIGTQVILNGARTDKDKLDTLPATIGTVLQGDYENFKSLTSLARKGYFD

Alignment Scores:
Pred. No.: 1.67e-17 Length: 10819
Score: 433.50 Matches: 164
Percent Similarity: 36.92% Conservative: 100
Best Local Similarity: 22.94% Mismatches: 262
Query Match: 9.17% Indels: 189
DB: 18 Gaps: 18

US-09-914-168-2 (1-919) x AE000493 (1-10819)
OY 211 AaNIlelyAlaIaleuclAspIletrGInglSerAlaMetaspleuSnglySer 230
DB 4941 AACGTTCTGCGCCAGCTTTCTACGATTGAAGTGAGTGACGCCACGCTGCTTT 5000
OY 231 lLeProArGleuArgGlnThrAlaLeuValAlaIaArgAlaValGlytyrTyr----- 248
DB 5001 CGCGACCGCGTCGATGATCCATCCCGCAAGGTCTGAAGCGCTGGATTATTAACGCGC 5060
OY 249 -----AspIleAspLeuSerIleIleArGAsnSerIleGlylValAlaSpaValIleIle 266
DB 5061 ACCATTGAATTGATCTCGCTCCACCGCAAGAGGGCGGCGAGTATTTGATTCGCCA 5120
OY 267 HIsAspLeuGlyGlnProValTyrIleAspTyrArGAlaValAlaGlnValArGlyGlyGly 286
DB 5121 GTACAGCGAGCGCTCGGTGTTAATTGGCGGACCGATGTTGCGGCGGCGGCGG 5180

```

```

OY 287 AlaSpAspLysAlaPheThrThrValAlaIaSpGlnValProLeuLeuIleGlyAspVal 306
DB 5181 CGACCGATTAAAGACTATTGAAATTCCTCGATCTCCGCCG---GCCATTGGACCGGTA 5237
OY 307 PheHIsHsGlyLysTyrGlnThrThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHIs 326
DB 5238 CTGAACCGAGCGCATTTATGAAATTTCAAAAATTCCTTAACACACATTTGCTCCGTAA 5297
OY 327 GLytyrPheAspGlyArGtyrPleuAspArgSerValAspValIleLeuProAspAsnThr 346
DB 5298 GGTATTTCGATACGGAATTATACCAAGCGCAGCTGGGCAATTCGGCTCGGCCCATAAA 5357
OY 347 AlaSpAlaSerLeuIleTyrAspThrGlyThrThrInlyrArGpheAspGlnValAlaPhe 366
DB 5358 GCCTTTCGGATTTGATTTATTAACAGTGGCAACGTTACGCTTGGGATGTGACCTTT 5417
OY 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386
DB 5418 -----GAGGATCACAATCCG 5435
OY 387 ArgGlnLeuLeuGlnGlnLeuLeuThrValAsnMetGlyGlnAlaTyrAsnLeuGlnAla 406
DB 5436 GATGAATACCTGCMAAATCTGGTCCGCTTAAAGAGGCGCATGATGCAATCGAAAGAT 5495
OY 407 ValArGAlaLeuSerAsnAspLeuIleAlaThrArGtyrPheAsnMetValAsnThrGlu 426
DB 5496 CTGGCAACAATGAACCCGCCACTTTCGTACCGGCTGTTTACTCGGTG----- 5546
OY 427 lLeValPheProGlnArGlnGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSer 446
DB 5547 GTGGTGCTCCACAA-----TTTGATTAAGCGCGC 5576
OY 447 SerSerArGThrGlnProAlaGlnValAspGlnSerThrLeuGlnProValIleGlnThr 466
DB 5577 GAACGCAATATTACCA----- 5594
OY 467 ValGlnLeuThrAspGlyIleLeuMetAspIleSerProIleGlnPheSerAlaSerAsn 486
DB 5595 -----TTGACG---GCGCG-----GTTTCGCGC----- 5615
OY 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArGHisLeuTyrArSpmetPro 506
DB 5615 ----- 5615
OY 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
DB 5615 ----- 5615
OY 527 lLeSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlnSerGluAsnGlu 546
DB 5616 -----CGAACAGAAACACCG 5630
OY 547 ValIleAspLeuProGlnArgThrAlaLeuAlaAsnArGlyThrProAlaAspValTyr 566
DB 5631 ATC----- 5633
OY 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArGAspGlyGln 586
DB 5634 -----GAA 5636
OY 587 lLeGlyLeuGlyTyrGlySerAspThrGlyThrArGLeuValThrLysPheGlnHisAsn 606
DB 5637 ACCGGGCTCGGTACTCTACGACGAGCGGAGCGCGCGTGAAGCGACGCGAAGAAAGCGC 5696
OY 607 LeuIleAsnArGAspGlyTyrGlnIaGlyAlaGlyLeuArGLeuSerLysLysLys 626
DB 5697 TGGATGAACCTTTATGGTACAGCTGTACACCCAGCAGCAGTACTAGTATTCGCGCGGCAAG 5756
OY 627 GLyValLysLeu---TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArG 645
DB 5757 ACCCTCGACTTGACGATTAATAATGCGCGTGTGAGAGAAATCCACTGCAACAATATTATTGG 5816

```

QY 646 AAlarhrleuglytyrclnglncgluValphegylHisSerThrAsnGlyPheAspSerLeuSer 665  
 Db 5817 GTGCAAGGGGGGTTTAAAGGCACT-----GACCTGAACGATACCGAATCTGAC 5864  
 QY 666 ThrAgtThrleuglHisgluIleSerArgSerIleIleGlnAsnGlyLysPheAsnArg 685  
 Db 5865 TCCACTAGCGCTG-----GTGGCTTCTCGCTCTCGGATCTCTCGCGCGCTGGCGGCT 5918  
 QY 686 ThrTySerleuArgTyArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrp 705  
 Db 5919 GCCATTAACTCGCTCGGTGGAGTCTCGAC----- 5945  
 QY 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuLeu----- 723  
 Db 5946 -----CACTTTACTCAGGCTGAATTAACCAATTAACCAACGATGCTGTTTAT 5990  
 QY 724 AlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTy 743  
 Db 5991 CCTGGGCGATGATTAAGCCGACGCGCTTCTCGTGTGGCTGATGCGCAACCTGGGGCGAC 6050  
 QY 744 ArgGlnArgTySerleuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763  
 Db 6051 TCGCAACCGCTACTATGACACTACACCAACAGCGCTGGGGTTCAATGCTCATTTCTCC 6110  
 QY 764 IleAlaArgAlaGlyIleSerGlyValTySerPheGlyAspAsnAlaTyrglySerAsn 783  
 Db 6111 GTTTTCACGGCGCAGACACTGTGATCCGACACTGACACTGAC----- 6155  
 QY 784 ArgAlaHisGlnMetThrGlyLysIleGlnAlaGlyTyIleTrpSerAspAsnMetAsn 803  
 Db 6156 -----CATCGTTTGTATACACGCGCGACGCTGGCTGTGATGAACCGGTGATTCGAC 6209  
 QY 804 HisValProTyArgLeuArgPhePheAlaGlyLysGlnSerIleArgGlyTyAla 823  
 Db 6210 AAAGTACCGCGCGATCGCTGCTTCGCGCGGGCGGACCGACGATTCGTGCTCAAA 6269  
 QY 824 HisAspSerLeuSerProIleSerAspLysGlyTyTyLeuThrGlyLysGlnValAla 843  
 Db 6270 TACAATCTATCTCGTCCGAATATACGCAACGCTGATGAAGGCGCTCGAATGTGATA 6329  
 QY 844 ValGlyThrAlaGluTyArgAsnTyArgIlePheMetLysAspLeuArgLeuAlaValPheGly 863  
 Db 6330 ACCGATCGCTGGAATACAGTACACAGTACCGCAAAATGCTGGCGCGGCTGTTGTC 6389  
 QY 864 AspIleGlyAsnAlaTyArgLysGlyPheThrAsnAspThrLysIleGlyAlaGlyAla 883  
 Db 6390 GATAGTGGCGAAGCGGTAAAGCATATTCGCCGACGAGCATTTAAACCGTTACCGGGTTC 6449  
 QY 884 GlyValAlaGlyTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyVal 902  
 Db 6450 GCGCGCGCTGGGAATTCGCGCGGCACTGAATTCGCGTACCGGTTCGCG 6509  
 QY 903 ---LysGlnGluGlyAsnProIleLysLeuHisPhePheIleGly 916  
 Db 6510 GATAAAGACGAACACGCG-----TTACAGTTTACATCGGT 6545

RESULT 16  
 AE005654 11760 bp DNA linear BCT 21-MAR-2001  
 LOCUS AE005654  
 DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 273  
 of 290.  
 ACCESSION AE005654 AE005174  
 VERSION AE005654.1 GI:12519212  
 SOURCE  
 ORGANISM Escherichia coli O157:H7 EDL933.  
 Escherichia coli O157:H7 EDL933.  
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 Escherichia.  
 REFERENCE 1 (bases 1 to 11760)  
 AUTHORS Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,  
 Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,  
 Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,  
 Grobbeck,E.J., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,  
 Grobbeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamoukis,K.,

Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,  
 Welch,R.A. and Blattner,F.R.  
 Genome sequence of enterohaemorrhagic Escherichia coli O157:H7  
 Nature 409 (6819), 529-533 (2001)  
 21074935  
 MEDLINE  
 PUBMED  
 11206551  
 2 (bases 1 to 11760)  
 Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,  
 Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,  
 Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,  
 Grobbeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Shao,Y., Potamoukis,K.,  
 Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,  
 Welch,R.A. and Blattner,F.R.  
 Direct Submission  
 Submitted (22-OCT-2000) Laboratory of Genetics, University of  
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

FEATURES  
 source  
 1..11760  
 /organism="Escherichia coli O157:H7 EDL933"  
 /strain="EDL933"  
 /serotype="O157:H7"  
 /db\_xref="taxon:155864"  
 /note="enterohemorrhagic"  
 gene  
 126..866  
 /gene="cysQ"  
 /note="25825"  
 126..866  
 /gene="cysQ"  
 /note="25825"  
 126..866  
 /gene="cysQ"  
 /function="phenotype; Central intermediary metabolism:  
 sulfur metabolism"  
 /note="Residues 1 to 246 of 246 are 99.18 pct identical to  
 residues 1 to 246 of 246 from Escherichia coli K-12 strain  
 M61655: B4214"  
 /codon\_start=1  
 /transl\_table=11  
 /product="affects pool of  
 3'-phosphoadenosine-5'-phosphosulfate in pathway of  
 sulfate synthesis"  
 /protein\_id="AAG59412.1"  
 /db\_xref="GI:12519213"  
 /translation="MDQYCOLARNGDAIMQYDGAQKPMQDVSKADNSPVTAADIA  
 HVIYNDGLTLPDIPVLSSEDPGKVEYKQHWRYLYLPDQTKFKIRNGEFTNI  
 ALIDHKPILGVAVPVMNVMTSAAGKAKKECGVKQIOYNDAPPLVIVSRSHAD  
 ABLKEYLDQLGHEHTTIGSSLSKFLVAGQADLYPRFPNTIMDTAAGAAVAAAAG  
 HYHDMQKPLDPTPRESFLNPGFRVSIY"  
 1078..2016  
 /gene="25826"  
 1078..2016  
 /gene="25826"  
 /function="orf; Unknown function"  
 /note="Residues 1 to 133 of 312 are 98.49 pct identical to  
 residues 1 to 133 of 166 from Escherichia coli K-12 Strain  
 M61655: B4215"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="AAG59413.1"  
 /db\_xref="GI:12519214"  
 /translation="MLPRIRHNNEFGAVELFYKSSYTKTHSNFENNIIHAFKKDMMI  
 SNYDSLITRREFRCATQTDKSSYQVLSKNEVYVNMAMDFLFSFKDNGAETITLRL  
 GSGFEYETIRPTINEYSFMDENRREPLQONRRLTAWDILQKQSDIPRKIGYTHQ  
 AIGDVSILGALDDIDIVSKLRGFELOGGGCKQKODGWIYKIIDNFAIYESIG  
 FALKIIGCGERFNGSSVLEDEYDLINFLVNACQOQVEIAEQDAIVASNLADS  
 DITKERICKYKSTIEAFKKDEALPVLVLRKNSER"  
 complement(2079..2633)  
 /gene="yrfJ"  
 /note="25827"  
 complement(2079..2633)  
 /gene="yrfJ"  
 /function="orf; Unknown function"  
 /note="Residues 1 to 184 of 184 are 98.91 pct identical to  
 residues 1 to 184 of 184 from Escherichia coli K-12 Strain  
 M61655: B4216"  
 /codon\_start=1



OY	327	GATYrPhaspsolYlrgrtrpleuAsparSerValaspallleuProaspnrth	346
Db	6238	GGTATTTCGATAGCAATTTACCAAGCCAGTCGGCATTCGGCTCGCATAAA	6297
OY	347	AlAspvalSerleuIeYrAsprthrglyThrlnYrAphAspIgluValIphe	366
Db	6298	GCCTTCGGATATGATTATATACAGCGGAACGTACCGCTTGGGCATGTGACCTT	6357
OY	367	PhethrIleAspProlystrAsnglnleuThrThrAspProaspIysleuProvalLys	386
Db	6358	-----GAAGATCACAAATCCG	6375
OY	387	ArgIuLeuLeuGlnIuInleuLeuThrValAsnMetGlyIuAlaYrAsnleuGlnAla	406
Db	6376	GATGAATACCTGCAGAAATCTGGTCCGTTTAAAGAGCGCATGATCGAATCGAAAGAT	6435
OY	407	ValArGAlaLeuSerAsnSpleuIleAlaThrArGtyrPheAsnMetValAsnThrglu	426
Db	6436	CGGCGAAGACTGAACCGCCGACTTCTGCACCGCGCTTAACTCGGtg-----	6486
OY	427	IleValPheProGluArGlnGlnIleGlnAsnAspIuValSerPheGlnSerSer	446
Db	6487	GGGTGGCTCCACAA-----TTTGATTAAGCCGCG	6516
OY	447	SerSerArGThrGluProAlaGlnValAspIuSerThrLeuGluProValIleGluThr	466
Db	6517	GAACGCAAGATTACCA-----	6534
OY	467	ValGluLeuThrAspIleIleuMetAspIleSerProIleGluPheSerAlaSerAsn	486
Db	6535	-----TTACG--GGCGGg-----GTTTCGCGG-----	6555
OY	487	LeuIleGlnAspIysLeuAsnleuValAlaAlaLysAlaArGlnIleuTyrrAspMetPro	506
Db	6555	-----	6555
OY	507	AspAspArGValLeuAlaIleAsnHisAspAspGlyValAsnArGSerIleleuGlyArg	526
Db	6555	-----	6555
OY	527	IleSerAspAlaValSerAlaValAlaArGAlaIleleuProAspGluSerGluAsnGlu	546
Db	6556	-----CGAACGAAACACT	6570
OY	547	ValIleAspLeuProGluArGThrAlaLeuAlaAsnArGlyThrProAlaAspValTyrr	566
Db	6571	ATC-----	6573
OY	567	GlnSerLysLysValProleuTyrrValPheValAlaSerAspIysProAlaAspGlyGln	586
Db	6574	-----GAA	6576
OY	587	IleGlyLeuGlyTrpIySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsn	606
Db	6577	ACCGGGCGGTACTCTACGACGCGGGAACCGCGCTGAAGAGCAGCTGGAAAAAACCG	6636
OY	607	LeuIleAsnArGAspGlyTyrrGlnAlaGlyAlaGluLeuArGLeuSerGluAspIysLys	626
Db	6637	TGCATGATCACTCATACGCTCAAGCTGCACCAAGCAAGTACAGATATTCGGCCCGCAACAG	6696
OY	627	GlyValLysLeu---TyrrAlaThrLysProLeuSerHisProLeuAsnAspGlnleuArG	645
Db	6697	ACCCTCGACTTCGCTATATAATGCCCGCTCTGAAGATACCACTGGAACATATTATTGG	6756
OY	646	AlaThrLeuGlyTyrrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer	665
Db	6757	GTGCAGGGGGGTTTAAAGCGCACT-----GACCTGAACGATACCGACTTCGAC	6804
OY	666	ThrArGThrLeuLunHisGluIleSerArGSerIleIleGlnAsnGlyTyrrAsnArG	685
Db	6805	TCCACTACCGTG-----GTGGCTTCGCTACTGGCATCTCTCCACGGCTGGCAGGT	6858

OY	666	ThyrisSerLeuThrArgLeuLysPheLeuLysThrGlnAlaProProGluThrTrp	705
Db	6859	GCCATTAACTCGCTGGAGCTCGAC-----	6885
OY	706	GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeu-----	723
Db	6886	-----CACATTCTCAGGGTGAATAATACCAACACACGATGCTTTTAT	6930
OY	724	AlaGlyValAlaValAlaHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyLys	743
Db	6931	CTTGGGGATGATGATTAGCGCACGGCTCTCTGGTCGTGATCCCAACCTGGGGCGAC	6990
OY	744	ArgGlnArgTyrSerLeuGlnValAlaGlySerSerGlyLeuValSerAspAlaMetAla	763
Db	6991	TGCGAACGCTACTCTATGACTACTCCACACACGGCCTGGGGTTCAAGATGTGATTCTTC	7050
OY	764	IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn	783
Db	7051	GTTTTCACGGCGCAACAGCTGTGATCGCACACAGTACGATCGC-----	7095
OY	784	ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsn	803
Db	7096	-----CAWCGTTTGTGCACACGGCGCACCGTGGCTGATGAACCGGTGATTTCGAC	7149
OY	804	HisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAla	823
Db	7150	AAAGTACCGCCCGGATCGGTTTCTTCGCCGGGGGCGATCGCATTTCCGCGTTACAA	7209
OY	824	HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAla	843
Db	7210	TACAAATCATTCGCGCCCGGAATACGCTTAAGCTGATCATAAAGGGGCTCGAAGTTGATA	7269
OY	844	ValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly	863
Db	7270	ACCGGATCGCTCGAATACACAGTACACAGTGACCGGAATAATGCTGGGGCGGCTTTGTGC	7329
OY	864	AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyAla	883
Db	7330	GATAGTGGCGAAGCGGTAAAGCATTTTCGCCCGACGACTTTAAACCGGTACCGGGGTC	7389
OY	884	GlyValArgTyrPalasSerProValGlyGlnValArgValAspValAlaThrGlyVal---	902
Db	7390	GGCGTACCGTGGGATTCGGCGGTGGGGCCATCAACCTGATTTGGCGTACCGGTGGCG	7449
OY	903	---LysGluGluGlyAsnProIleLysIleHisPhePheIleGly	916
Db	7450	GATAAAGACGAACACGGG-----TTACAGTTTTCATACGTGT	7485
RESULT 17			
LOCUS	ECOW93	338534 bp	DNA linear BCT 30-JAN-2001
DEFINITION	Escherichia coli K-12 chromosomal region from 92.8 to 90.1 minutes.		
ACCESSION	U14003		
VERSION	U14003.1	GI:1263172	
KEYWORDS			
SOURCE	Escherichia coli.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
REFERENCE	1 (bases 1 to 338534)		
AUTHORS	Burkard,V., Plunkett,G. III, Sofia,H.J., Daniels,D.L. and Blatner,F.R.		
TITLE	Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes		
JOURNAL	Nucleic Acids Res.	23 (12),	2105-2119 (1995)
MEDLINE	95334362		
PUBMED	7610040		
REFERENCE	2 (bases 1 to 338534)		
AUTHORS	Plunkett,G.I.,I.I.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-AUG-1994)		
	University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.		
	Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:		

COMMENT	FEATURES
608-263-7459 On Apr 17, 1996 this sequence replaced gi:536929. This sequence was determined as part of the E. coli Genome Project (Frederick R. Blattner, director) at the University of Wisconsin-Madison. Supported by award H600301 from the NIH Human Genome Project. The entire sequence was independently determined from E. coli MG1655; overlaps with other sequence determinations are annotated. The start of this entry overlaps the end of the entry ECOW089 (000006) by 1885 bp.	Source Location/Qualifiers 1. 38534 /organism="Escherichia coli" /strain="K-12" /sub_strain="MG1655" /db_xref="taxon:562" /map="92.8 to 100 minutes" /note="This sequence comprises the following lambda clones: DD928(EC27-278), DD930(EC30-188), DD933(EC27-1320), DD935(EC30MM2), DD937(EC27-1070), DD941(EC17-142), DD945(EC17-8), DD947(EC34A-34), DD949(EC22-169), DD952(EC27-297), DD953(EC27-409), DD956(EC17-103), DD958(EC30MM1), DD960(EC30MM2), DD962(EC21-104), DD965(EC22A-40), DD968(EC30K60A-4pp), DD970(EC19-202), DD974(EC19-61), DD975(EC18-233), DD977(EC30MM7), DD980(EC27-1151), DD981(EC18-3), DD984(EC17-101), DD987(EC18-115), DD990(EC17-136), DD992(EC18-282), DD995(EC17-115), DD997(EC18-113); M13mp19 or Janus vectors were used for subcloning" 1. 1885 /note="1885 bp overlap with end of GenBank Accession Number U00006 (ECOW089)" Complement(<1. 398) /note="ORF_f336 of GenBank Accession Number U00006 (ECOW089)" /codon_start=1 /transl_table=1 /label=ORF_f332p /protein_id="AA06985.1" /db_xref="GI:1263173" /translation="MGFTTRVKSASEKRPFNALFMDKYGTFITLIIAIFGSLSP EYELTNINIOIFVOSSVTVLIGGEFFALLVAGIDSLVGAILALSGMTAKMLAGV DPFLAAMIGVVALGVALGALINGCLVNTGL" 277. 397 /note="predicted bend of 75.10 degrees" Complement(377. 1909) /note="ORF_f510" /codon_start=1 /transl_table=1 /label=ORF_f510 /protein_id="AA06986.1" /db_xref="GI:536931" /translation="MATPTISMAGICGRSEPPVAAKLSVNLTVYPGEITHALLGENGAK STLMKLVLSGHEPTKCTITTNISYNKLDHRLAAQLGIGITVYELSVIDELVLENLV IGHRLTKIGCVNIIDWRKMRVRAAMMLRVGLKVDLDEKVAANLSISKMLLEAKTL GMLAKVIMDEPTSLTKEVDYLRFLIMNOLRKEGTAIVYISKLAIEIRICRYTYM KDSSVCSGIVSDVSNNDIIVRLWGLRQLEFRANEMKENVNLHETVEFVRYNTSRBR AKYRDLSFVSQREGIILGPAVLGSGREGLNCLGVQDKRAGCGTILNKGKISPSPLD AAKCKMAYITRESRDNGFPNPSIAOMALISRLKDGCGAGLREHVEDQRAENQ RELALKCHSVNONITELSGNOOKVLISMLCCCPVYIIFDEPTRKIDVGAAEAYK VMQQLADDGKVIILMWSELPEIITVCDRIAVFCGRITLTILNRDMSDEEIMAWALP QE" 809. 931 /note="predicted bend of 74.75 degrees" 1768. 1889 /note="predicted bend of 75.87 degrees" Complement(1983. 2010) /note="promoter matrix score of 55; putative; resembles flh-regulated promoter" Complement(1985. 2021) /standard_name="REP; repetitive extragenic palindromic element" /note="contains 1 REP sequence" Complement(2036. 2971)

misc\_structure

CDS

/note="ORF\_f311"

/codon\_start=1

/transl\_table=11

/label="ORF\_f311"

/protein\_id="AA96987.1"

/db\_xref="GI:536932"

/translation="MNRKYKYPGSGTLVGLMLSTAPFAAEYAVVLKTLSPNPFWMQMKR  
GJEDKEKIVASVDIFASPSEGDPOSOLQEPDLSENKRYKGIAFPILSSVNLVDPVA  
AMKKGYIYLNDKIDMDLMLKAGGNVEAFYTDDVAYAKGASTITDKLAGEGEVA  
IEBKGNGNSGERNRGATEARFKASQIKLVASPADWDRIKALDVATNVLRNNIN  
AITYCADMTVAEQAQAVANAGTGRVLVVGTDGIEPARKMEAGOMTATVANPADIG  
ANGLKMTVAEKSGKVIPLDKAPEFKLVDLSILVTG" \*

3010 . 3121

/note="predicted bend of 73.22 degrees"

complement(3030..3953)

/note="ORF\_f307"

/codon\_start=1

/transl\_table=11

/label="ORF\_f307"

/protein\_id="AAA96988.1"

/db\_xref="GI:536933"

/translation="MASSIOGEGENVMSOSEPSALPNIGLAPLYRMKOEQTENES  
IYEMLEPKNTSCAPAIRDVAALVALVSAMLYKSKLLGFSGFRMLRSALDITQSOS  
QVLPSELADEAPQDVNKVENFTILTITMEGSTLVNDEIHRAAFYLOAQROLYGA  
GSNAICCAVOHQRFLRIGRCQAYPDAMIHMNASLALOEGVAVLVHTSGTSIDVKAA  
VBLNMGNGRKIICTIHSHSYHSPIAKLADYITICSPEPEPLDGRMASRIQLTLDAAFE  
VESAOINTEOANIIMQKGIADVDFESPALK"

complement(3952..3979)

/note="promoter matrix score of 61. putative"

4151..4251

misc\_structure

gene

/note="predicted bend of 72.06 degrees"

4279..4728

/gene="yjcA"

4279..4728

/gene="yjcA"

/codon\_start=1

/transl\_table=11

/label="ORF\_0149"

/protein\_id="AA96989.1"

/db\_xref="GI:536934"

/translation="MKKIAGCGDHVGFILKHIEAHVLNERGVEVIDKGTMSSETDPP  
HYASOVALLAAGEVDGGLICGTGVGISIANKFAGIRAVVCSEPYSAQISRQHNDT  
NYLAGSRVGLGELAKMLIVDMVLGAQYEGGRHQREVAITALEQRN"

complement(4520..16191)

/note="corresponds to GenBank Accession Number D90227  
(ECOPHN) 1..11672; the related sequence GenBank Accession  
Number J05260 (ECOPHNAQ) is from E. coli B, not K-12"

complement(4733..5146)

/gene="phnQ"

complement(4739..5146)

/gene="phnQ"

/note="annotated in GenBank Accession Number J05260  
(ECOPHNAQ), but probably not a gene"

/codon\_start=1

/transl\_table=11

/protein\_id="AAA96990.1"

/db\_xref="GI:536935"

/translation="MPAPYTPSSPLLIRHSPFLIQPLTLTIKLALAIYDPAIFVL  
AAAVVILSPAHEFTLPAVVARIPTVASPTVRVSVLNRLINTLANGVNINDIGLRAND  
SVSRVKNRNSGRFTISPOKETSLITLIITHPC"

complement(5272..6030)

/gene="phnP"

complement(5272..6030)

/gene="phnP"

/codon\_start=1

/transl\_table=11

/protein\_id="AAA96991.1"

/db\_xref="GI:536936"

/translation="MSLTITLTGTGAQGVPMWGCEACACARAPSQVRRQCSGVV  
KFNARITIDAGLHDLADRMRPGSFQDFLITHYHMDVQGLFLPRMGVGDPITPYEGP  
DEQGCDLIPKRPGLDSSHIVEPVFDLQGLQVTPRLPHNSKTLEYILLETAHSRYA  
WLTHSDHPPEKTKLFELRNNOQVMVAMDCSHPPRADARNCHDNLTVIALNQVIRSPRV





```

QY 724 ALAclYValAlaValAlaHlsYThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
    ||||| : : : ||| : : : |||
Db 134374 CCGGGGATGATGATTAGCCGACCGGCTTCGTGGGCTCATGCAACCTGGGCGAC 134433
QY 744 ARGcInaGTYrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
    ||||| : : : ||| : : : |||
Db 134434 TCCGAACCTACTCTATGACACTCACAACACGCGCTGGGCTGAGATGTCATCTCTCC 134493
QY 764 lLeAlaArgAlaGlyLlSerGlyValTYrSerPheGlyAspAsnAlaTYrGlySerAsn 783
    ||||| : : : ||| : : : |||
Db 134494 GTTTCACACGCCACAGACTCTGGATCCGACACTGACGATCC----- 134538
QY 784 ARGAlaHlsGlmethrGlyGlylleglnAlaGlyTYrFlerPserAspAsnPheAsn 803
    ||||| : : : ||| : : : |||
Db 134539 -----CATCGTTTGTTAACACGCGCACGCTGGCTGATTGAACCGGATTTTCGAC 134592
QY 804 HlsValProTYrArgLeuArpPheAlaGlyGlyAspGlnSerLleArgGlyTyrAla 823
    ||||| : : : ||| : : : |||
Db 134593 AAGTACCGCGGATCTCGTTCTTCGCGGGGCGACCGCAGATATTCGTGGCTACAA 134652
QY 824 HlsAspSerLeuSerProLlSerAspLysGlyTYrLeuThrGlyGlnValLeuAla 843
    ||||| : : : ||| : : : |||
Db 134653 TACAATATCTATTCGCTCCGAATATACCCAAACGCTGACCTGAAGGGCCCTCGAAGTTGATA 134712
QY 844 ValGlyThrAlaGlyTYrAsnTYrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
    ||||| : : : ||| : : : |||
Db 134713 ACCGATCGCTCGATACAGTACAGTACACGCGAATAATGCTGGCGCGCTGTTTGC 134772
QY 864 AsPlleglYAsnAlaTYrAspLysGlyPheThrAsnAspThrLyslleglYAlaGlyVal 883
    ||||| : : : ||| : : : |||
Db 134773 GATAGTGGCGCAAGCGGTAAGCATATTGCGCGCAGCAGATTTAAACCGGTACCGGGGTC 134832
QY 884 GLYValArGTTPAlaSerProValGlyGlnValArgValAspValAlaThrGlyVal--- 902
    ||||| : : : ||| : : : |||
Db 134833 GCGCGCGCTGGGATCGCGGCTGGCCAAATCAATCGATTTGGCTGACCGGTCGCG 134892
QY 903 ---LysGluGluGlyAsnProLlSerLeuHlsPheMetllegly 916
    ||||| : : : ||| : : : |||
Db 134893 GATAAAGCAACACGCG-----TTACGATTTTACATCGGT 134928

RESULT 18
AX033470 1731 bp DNA linear PAT 21-SEP-2000
LOCUS AX033470
DEFINITION Sequence 3 from Patent W00047737.
ACCESSION AX033470
VERSION AX033470.1 GI:10280230
KEYWORDS
SOURCE
    Haemophilus influenzae.
    Haemophilus influenzae
    ORGANISM
        Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
        Haemophilus.
REFERENCE
    1 (bases 1 to 1731)
    Ruelle,J.L. and Thonnard,J.
    TITLE
        Haemophilus influenzae rd outer membrane sequences used as vaccine
        JOURNAL
            RUELLE JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ; THONNARD
            JOELLE (BE)
FEATURES
    source
        1. 1731
        /organism="Haemophilus influenzae"
        /db_xref="taxon:727"
BASE COUNT 509 a 322 c 386 g 514 t
ORIGIN

```

Alignment Scores:

Pred. No.:	Score:	Length:
3.56e-18	427.50	1731
Percent Similarity:	35.64%	Matches:
Best Local Similarity:	20.35%	Conservative:
Query Match:	9.04%	Mismatches:
		Indels:
		Gaps:

US-09-914-168-2 (1-919) x AX033470 (1-1731)

```

QY 243 ARGAlaValGlyTYrTYrAspLleAspLeuSerLleArgAsnSer---lleglyGlu 261
    ||||| : : : ||| : : : |||
Db 211 CGTGCTTTGGTTATATGATCTCCGCTTTGAAGAAACGCTCAAGCGCAA 270
QY 262 ValAspValLlelLeHlsAspLeu-----GlyGluProValTYrLleAspTYrArgAla 279
    ||||| : : : ||| : : : |||
Db 271 CCGGATTTATGATTCCTCATGTTACACCGCGGACCAACAAATTCGGGGAGCAT 330
QY 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluVal 299
    ||||| : : : ||| : : : |||
Db 331 GTGCAATTCAGGGGAGAGCGCAAGATGAATAATTTGATCGCTAATAAACTTG 390
QY 300 ProLeuLeuLlleglYAspValPheHlsHlsGlyTYrGlyThrLysAsnLeuLle 319
    ||||| : : : ||| : : : |||
Db 391 CCA---AAGAGAGCGCTTTGGTTGGAACCAACTTACATATTCACAAACGCAAT 447
QY 320 GluAsnAlaSerAlaGluHlsGlyTYrPheAspGlyArgTYrLeuAspArgSerValAsp 339
    ||||| : : : ||| : : : |||
Db 448 TCACGCTTGCAATTAATCGTGATTTGATGGAACTTTAAATTCACGTTAGAA 507
QY 340 ValLleLeuProAspAsnThrAlaAspValSerLeuLleTYrAspThrGlyThrGlnTYr 359
    ||||| : : : ||| : : : |||
Db 508 ATCAGCCCTGAACCCCATCAAGCATGCGGCAATGTATTTATAGTGCTCCGTTAT 567
QY 360 ArgPheAspGluValAlaPhePheThrLleAspProLysThrAsnGlnLeuThrAsp 379
    ||||| : : : ||| : : : |||
Db 568 CATTAATGCAATATTAATCTTT----- 588
QY 380 ProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGly 399
    ||||| : : : ||| : : : |||
Db 589 ---AGCCATTTCACAAATTCGCTGACGATTAATCTCAATAAATTTCTTAACATTCGCG 645
QY 400 GluAlaTYrAsnLeuGlnAlaValArgAlaLeuSerAspSerLeuAlaThrArgTYr 419
    ||||| : : : ||| : : : |||
Db 646 GATCCATATTATTAATTAATTTGCGATTTAACACGCAATTTTCATCTTAATTTGG 705
QY 420 PheAsnMetValAsnThrGluLleValPheProGluArgGluGlnlleglnAsnAspGln 439
    ||||| : : : ||| : : : |||
Db 706 TTTAGCTCAGTA----- 717
QY 440 ValSerPheGluGlnSerSerSerArgTYrGluProAlaGlnValAspGluSerThr 459
    ||||| : : : ||| : : : |||
Db 717 ----- 717
QY 460 LeuGluProValLlleglunThrValGluLeuThrAspGlyTYrLeuMetAspLleSerPro 479
    ||||| : : : ||| : : : |||
Db 717 ----- 717
QY 480 lleglunPheSerAlaSerAsnLeuLlleglnAspLysLeuAsnLeuValAlaAlaValAla 499
    ||||| : : : ||| : : : |||
Db 718 -----TTAGTTACGCTTAATGTTAAT----- 738
QY 500 ArgHlsLeuTYrAspMetProAspAspArgValLeuAlaLleAsnHlsAspAspGlyVal 519
    ||||| : : : ||| : : : |||
Db 738 ----- 738
QY 520 AsnArgSerLleLeuGlyArgLlSerAspAlaValSerAlaValAlaArgAlaLleLeu 539
    ||||| : : : ||| : : : |||
Db 738 ----- 738
QY 540 ProAspGluSerGluAsnGluValLleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
    ||||| : : : ||| : : : |||
Db 738 ----- 738
QY 560 LysThrProAlaAspValTYrGlnSerLysLysValProLeuTYrValPheValAlaSer 579
    ||||| : : : ||| : : : |||
Db 739 -----CATAAAGCAAAACCTGGATGAGATGAGATTTATCTTTATCCA 780
QY 580 AspLysProArgAspGlyGlnlleglYLeuGlyTYrPGlySerAspPheThrArgLeu 599
    ||||| : : : ||| : : : |||
Db 781 CGTAAATAAATGCGATGACCTCGGTGCGCTTGTCTACTGATGCGCGCTTACGGA 840

```

```

Oy 600 ValThrLysPheGluHisAsnLeuLeuIleAsnArgAspGlyTyrGlnAlaGluLeu 619
    |||||
Db 841 CAATATGGCTGGCAAAACCTTGATTAATAGCCGTGGACATAGTTGGCTTAATCTT 900
Oy 620 ArgLeuSerGluAspLysLysGlyValLys---LeuTyrAlaThrLysProLeuSerHis 638
    |||||
Db 901 TATCTCTCTGCACAAACAAACACTAGAGGCAACTTATCGAATGCGACTGTTAAATAAT 960
Oy 639 ProLeuAsnAspGlnLeuArgAlaThrLysLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
    |||||
Db 961 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1008
Oy 659 ThrAsnGlyPheAspLeuSerThrArgThrLysGluHisGluLysSerArgSerIleIle 678
    |||||
Db 1009 GAGAAC-----GATCCATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1062
Oy 679 GlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLysThr 698
    |||||
Db 1063 AATGCCCATGCTGGCAAAATTTTGGCGGACCTTCGTAACGCGATACGACAGCTTTT---ACA 1119
Oy 699 GlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSer 718
    |||||
Db 1120 CAACGGGAT----- 1128
Oy 719 GlnGluAlaLeuLeuAlaGlyValAlaValAlaHisLysThrValAlaAspAsnLeuValAsn 738
    |||||
Db 1129 -----ATCACATGATTAACCTTA-----CTTCTTTAT 1155
Oy 739 ProMetArgGlyTyrArg----- 744
    |||||
Db 1156 CCAACTGTGGATGATTTACGTCGACATGATACGTCGTGTTCTTTCGCACTTGGGGCGAT 1215
Oy 745 ---GlnArgTyrSerLeuGluValGlyLysSerGlyLeuValSerAspAlaAsnMetAla 763
    |||||
Db 1216 GTGCCAAAATTAATCTTTGATTTAAGCAAAAGCAATTTGGCTATCAGATCTTTTATA 1275
Oy 764 IleAlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn 783
    |||||
Db 1276 AAGTGCAGCAATCTGACGCGTGGCTGATATGACAGAAAT----- 1320
Oy 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsn 803
    |||||
Db 1321 -----CATCGTATCGTCTGCTCGTGAATGCGGATTTTATCAAAAGATATTGAA 1374
Oy 804 HisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAla 823
    |||||
Db 1375 AAAATTCGCCCTACACTGCGTCTTTCGTCGTGGCGATCGTAGTGGCGGTTACGCG 1434
Oy 824 HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyValAla 843
    |||||
Db 1435 TATTAATAAATTCGCCCTAATAATGAAATTTGGTGGCGCTCACGCTTCT 1494
Oy 844 ValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
    |||||
Db 1495 ACCACTTCTTAGAAATCAATATCAAGTTTATCAAAATGTCGGCGCACTTTTGA 1554
Oy 864 AspIleGlyAspAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883
    |||||
Db 1555 GATAGTGGATTAGCTCGCATTAATACACAGCAAAAGACCTGCTTATGGCACAGGCGTT 1614
Oy 884 GlyValArgTyrPheLeuSerProValGlyGlnValArgValAspValAlaThrGlyValLys 903
    |||||
Db 1615 GGTGTGCTGGGATCGCCAGTGGGATTAATTTGATATTCGACACACCATTCGT 1674
Oy 904 GluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
    |||||
Db 1675 GATTAAGATTAACAGCAAAATATTAATTTTACATCGGA 1713

```

```

RESULT 19
AE012531/ 11870 bp DNA linear BCT 23-MAY-2002
LOCUS Xanthomonas campestris pv. campestris str. ATCC 33913, section 439
DEFINITION of 460 of the complete genome.

```

```

ACCESSION AE012531 AE008922
VERSION AE012531.1 GI:21115353
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 11870)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardoso,J.,
Camargo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locall,E.C., Machado,M.A.,
Machado,M.A., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spindola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.
Comparison of the genomes of two Xanthomonas pathogens with
differing host specificities
Nature 417 (6887), 459-463 (2002)
JOURNAL MEDLINE
PUBMED 22022145
REFERENCE 12024217
AUTHORS
2 (bases 1 to 11870)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardoso,J.,
Camargo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locall,E.C., Machado,M.A.,
Machado,M.A., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spindola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.
Direct Submission
Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de
Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
Brazil
FEATURES
source
location/Qualifiers
1..11870
/organism="Xanthomonas campestris pv. campestris str. ATCC
33913"
/strain="ATCC 33913"
/db_xref="ATCC:33913"
/db_xref="taxon:190485"
/note="pathovar: campestris"
480..2099
/gene="XCC4080"
480..2099
/note="XCC4080"
480..2099
/gene="XCC4080"
/note="identified by sequence similarity; putative: ORF
located using Blastx/Glimmer/genemark"
/codon_start=1
/transl_table=11
/product="polyvinylalcohol dehydrogenase"
/protein_id="AA043301.1"
/db_xref="GI:21115354"
/translation="MSIDPRAFAASQIWRFSQSNLPFASPNMWPAGGYSFNSRA
PGFWKINRTTAPLKPAMTITTTGDISATFSVBGGLTYVPDMGCTLYSDTOSKRPW
QARLSDTGNANSVTRNTPAISGKSIIVGQASGTVADIKTKGKLWSSVAESNPGA
RITSBPVIYGNRIYVIGISGDMGNLANHTFSFGSVVALDLTKKKVQFVNVPEGY

```

TCASVWGVAIDPISRLRYFGTCNNYNIIPAGVNCVSNARSTRGSELTIVODELNCMA  
DNAYVSVALNLDTGKLMADAAQGYDAMNLSCIVAPITNGICPNTQATNLTPDVEDG  
GAGVQLFTVMRDGKPOQLVAGACQSGIYAPARTGAKWSTQVPGGTTGCGEAGTA  
FDPYKQVYVAIINNSSNATYTLGPDNTVSHNGSNVADLAATGKIKMOVKVGVSVN  
PNVPAGGGSLTVSPNLLYAGSMACNMVALDPTDGTALMNYNATGSVSISSPAVVGCAL  
YWGAGYGRPNRTGATSNOLIKFVPSCAK"  
/complement(2282..2387)  
/gene="XCC4081"  
/complement(2282..2587)  
/gene="XCC4081"  
/note="putative: ORF located using Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA043302.1"  
/db\_xref="GI:21115355"  
/translation="MAIARRVICLGRSRMNIMAAARRPRAMSTLAPEROLLHYPADYE  
GYCDRPLVMGQLASVEYPAADVDAELVYALRRRAEAHGLEAGQIRLCQVRL"  
/complement(2490..2930)  
/gene="arod"  
/note="XCC4082"  
/complement(2490..2930)  
/gene="arod"  
/note="identified by sequence similarity: putative: ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="3-dehydroquinate dehydratase"  
/protein\_id="AA043303.1"  
/db\_xref="GI:21115356"  
/translation="MILGRPEASSLVGRPOLPRAVITTOLYRCARDAGKTYALRACG  
SEBELDLRVAGQARMEITALIDPSCVDVSHLVHVLGGLAIPVEAHDDSTDRPERC  
LPAGIGORVAIVRGCAQSYLGLIEALHLCGTETIGDHYGT"  
/complement(2950..3372)  
/gene="XCC4083"  
/complement(2950..3372)  
/gene="XCC4083"  
/note="putative: ORF located using Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA043304.1"  
/db\_xref="GI:21115357"  
/translation="MCILILRGPSDAIPVAPLPLCAGRDLTACADLASLIAGLQA  
APLAVITPHDAPRAYAMSLGIAACPLALHLMRYAA"  
/complement(3616..7572)  
/gene="XCC4084"  
/complement(3616..7572)  
/gene="XCC4084"  
/note="identified by sequence similarity: putative: ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA043305.1"  
/db\_xref="GI:21115358"  
/translation="MALGASAGGYCARLERPRFPAVALHRRNGEPVNTTPTSPA  
PRPRYRRRRFVSGGLVGLGLALALALVLAQVAVTLARLPVGSFT  
YGSVEYPAAGPLTLNVDPRYODIHFTAEVYLEEDLVRGRKLOLAVOVSNNTL  
LGKSDPEPTLPSPESLPIQNLVPLAIOADHIGVMDLRITLOOQPIITLHROGGLI  
TGELTRNLVYDTMGRPLRNDVYPSDEYRADLITRATVLAARGRTPASLGLVARGD  
LDKMEVAIAGRAPALRASLVFTGKDDPTMAFNATEALDLSLIIPAGGDETPPATP  
IALDLQASGKADLRGVSVOGELSATLQPSHLSLOQVLTVEPLIIDAEGRTPLR  
GSADEFDADNARSFRSVNASGLRTPADPATPDAPAVPELTDRGLLAKMAA  
IGRATVROGOAELVPSRGNDRALQKQVAKTPGSLDTGAVMAAPLOMLTA  
OLAKFPCYFVPGWGNLSGOVASKRQLPAPAGVSGCYENTADIPITLQGLRAL  
SANKFALRGTQEGQLDALGSSRTDKKTYGGDLDIAOLQPLQDLDLGAGRTV  
RGQLQVNRDAPVTADLACTGLKMWNYSAQSVSLRKLPMSSGDQLALQGTALIEA  
GVLIDSVNRQAGVAELRLDADIANMSVAVLQGDVRRNARNGOVSTLRIYPAK  
DAMALROPACSTEGAAFTLSDTCLGATGALCAGMPEGMVHAGCALSLVSPQ  
WLPROGROIYLRGLSLDGSFKPRGNMWSFRLASSEGVCRLGENTYAVAGPNR  
GELLRIYDQFSVQADFTTPOQIQGKLCVGTQGNCFDPAKINTGMDAYAPLNGELYLMSR

LYWLELVADVVRPKGLVEGHVSLRGTDRKPLLGADATLSDFEAYPSMGLTLSEKGC  
REDLALPDGSAKSTASAKSGDGLINIDGISTWGTSTPLLINTIRGENVLYNNTSELTIV  
ANDPQFEGIVDTTMOLRGVTYVPEADIDLEDRGTSVEDVYVDPVPEAPASPL  
DMDLAVLGDCKMNSGEGKGLSGGOMQIRARPRGMTANGDIDVGRYKAYGODLTI  
SRGQLTMNNNTYSDPRVSRARERKTDTAGIDVGRASPRADVWSPEAKSGSESV  
YLVGLGUSTASSDERTQOVSAASALASGSLIASOTIGKGLIDEGVESSTQSVLS  
GFGKYLSPKLYVGVSVSWGSSVTLTKLLSRGPDVEASSTVETKGSIMNREK"  
/complement(7470..9410)  
/gene="XCC4085"  
/complement(7470..9410)  
/gene="XCC4085"  
/note="identified by sequence similarity: putative: ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA043306.1"  
/db\_xref="GI:21115359"  
/translation="MPHVRPRDLPCEDLCMYVHDVAYVTHGALSTPPRRPNRS  
RSSRYPAPMRFLIRVAPVPLLCITLAGSFARGTIDKIEIKGLDGDDEMEINIEVS  
LSLYEAVKEGCEESRLEYLLAOERQTRREALPFGYSPETIDIAAPRDESLTVITV  
NNGEPYVRKOSHISTIGMAOERYLODOKOPEPRGQVESHROYASVRLTRAE  
RGYFPADFQORVAVTRAEMADIDILNMGSGRRYMGTYRFPDYDFRPLFELYWGL  
EGSYHKEGLDLRESLTRKDYFTTIDIDPKPEADQGRVYVDKTLPAKRTVYTS  
LSYSGESGAGVGRGVRRYVNSRGRKMDTQDLOAQRKSLTTSYRVPATRMWDGYTA  
SARLYDEQTEYIDLNRVKLITGSRSGOINRWSAIASINALRBRMFSGSDPTDAYE  
TSTLYPOLQANVYVNDRLFRPSGCOMFIRGAEAGSDPTNFGOLGOLRWFAGA  
GNGRYILREGEQTTWTSDLVAMPSPSLREFAGAPNSIRRYARREYVPPRPKDFEFA  
AKNYVYASAEYHLYKSGPMGGAIVFDSSASARDPDMWTGTGFGLRMWSPIVGPVAVD  
IAHGLNDPDSOFOLYDITCANL"  
/complement(9598..11694)  
/gene="g1ys"  
/note="XCC4086: g1ys(B)"  
/complement(9598..11694)  
/gene="g1ys"  
/note="identified by sequence similarity: putative: ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="g1ycyl-tRNA synthetase beta chain"  
/protein\_id="AA043307.1"  
/db\_xref="GI:21115360"  
/translation="MSEQPLILIECTEELPYKALPGLQAQAFEDGYLAGEKRAVAT  
RKDAKPLSTPRRLAVLLPGVATEQEPORSEVIGPYLUNLUDNEGKTRALACPAAG  
IDMTALERTSDKGEFVRHAYVPGAQAALPEILIREIAAMPTEKPRMGAHEYAF  
ARPVOMLVLFQDITVYIPAEILGVRGDRITRGRFMDGIDIALAAGDYIDALRAHVL  
VDADRARARIVEVDAAARQAGSARISDNEQVNVNLEVSAGVYCSFEERFLAVPO  
EALITEMEINOKFPPVLDGGRKLTBOFTIGIANIVSDEVAKGVYRVRPPADAKF  
FPDEDLKGLGEMAGCAGIASVYTOAKLTGTADYVARAALAEIAAOVGDAPPOARAA  
ELAKNDLOSRYNNEPELOGIAGRIHAKAAGEPSETSLAITIDAYODRPPGDDITASPL  
GKVLAIARLDTLAGGFAGLKPNGKDPFALRRNALGLARTVIESGFDLPLKLDV

## Alignment Scores:

Pred. No.:	8,56e-14	Length:	11870
Score:	375.50	Matches:	161
Percent Similarity:	35.298	Conservative:	86
Best Local Similarity:	23.008	Mismatches:	232
Query Match:	7.948	Indels:	201
DB:	1	Gaps:	21

US-09-914-168-2 (1-919) x AE012531 (1-11870)

QY	235	ArgGlnThrAlaLeuValAlaAlaArgAlaValAGLYTYTYrAsP-----ILeAsPleu	252
DB	9032	CGGACAGACCGGC---GAGGCACTGGAGCGGTTGATTATTCAGCGCCACGACATCGACATC	8976
QY	253	SeTlleIleArgAsnSerIleGlyIuValAspValIleIleHisAspleuGlyIuPno	272
DB	8975	GGCGGCGCGCGGACGAGCGGACCGCTGACCGTGATCATACGGTGAACCGTGGCGGCGG	8916
QY	273	ValTYIleAspYrIArgAlaValGIuValArgGIuGluGIuValAspAspIuAlaPne	292
DB	8915	GTCGGGTCGCGGACGTCATATTTCGATCACCGGCTGGGCGGACGACGACGCTATCTG	8856

QY 293 ThrThrValAlaAspGluValProLeuLeuIleGluValAspValPheHisGlyLysTyr 312  
 DB 8855 GGGCAGGACTCAAGCGATTGCGAGCCGCGAGGCGAGGCTCTTCAGCCATCCACAGTAC 8796  
 QY 313 GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArg 332  
 DB 8795 GAAGCCAGCAGGTGGCGGCTACACCGCGCGCTCGGCGAGCCGCTATTTCGAGCCGAC 8736  
 QY 333 TrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIle 352  
 DB 8735 TTCACCCAGCGCGCGGCGGTGACCGCGTCCGACGATCGCGCGCATGATCTGAC 8676  
 QY 353 TyrAspThrGlyThrGlnTyr-----ArgPheAspGluValValPhePhe 367  
 DB 8675 TGGGACAGCGCGCGCGCTCGACATGCGACATGCGCGCTGCTGACATGATTTCTC-- 8619  
 QY 368 ThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArg 387  
 DB 8618 -----CGCCCG 8613  
 QY 388 GluLeuLeuGluGlnLeuLeuThrValAlaAsnMetGlyGluAlaTyrAsnLeuGlnAlaVal 407  
 DB 8612 GGGCTGTTCAGCGCGCTGATTGGAGCGGCGACGCTACTACACGAGGAGGAGCGTGC 8553  
 QY 408 ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIle 427  
 DB 8552 GATCGGCTGGCGGCGAGTACGACGACGACGATGCTTACCTACACATCGACATCGAC-- 8496  
 QY 428 ValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSer 447  
 DB 8496 ----- 8496  
 QY 448 SerArgThrGluProAlaGlnValAlaAspGluSerThrLeuGluProValIleGluThrVal 467  
 DB 8495 -----CCCAAGCGCGAGGAGGAGCCGATGACGAGCGCGCGTGC-----GACCTC 8448  
 QY 468 GluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu 487  
 DB 8447 AACTGAC----- 8439  
 QY 488 IleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp 507  
 DB 8439 ----- 8439  
 QY 508 AspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIle 527  
 DB 8439 ----- 8439  
 QY 528 SerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluVal 547  
 DB 8438 -----CGCGCC----- 8433  
 QY 548 IleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGln 567  
 DB 8432 -----AAGCCGAC-----GTGATGAC 8415  
 QY 568 SerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIle 587  
 DB 8414 TCC----- 8412  
 QY 588 GlyLeuGlyTyrPheLysSerThrGlyThrArgLeuValThrLysPheGluHisAsnLeu 607  
 DB 8411 GGTCTGAGCTATGCGACGAGGCGCGGTGCGCGCGGCGCGGCGGCGGCGCTAT 8352  
 QY 608 IleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGly 627  
 DB 8351 GTCAATTGCGCGCGCGAGATGACACGACGCTGACTACGCGAGAACCGCAAGAGC 8292  
 QY 628 ValLysLeuTyrAlaThrLysProLeuSerThrProLeuAsnAspGlnLeuArgAlaThr 647  
 DB 8291 CTGACACACAGTACCGCGGTGCGCGCTTCGCGTGGATGGTGTGATACCGCGCTCG 8232

QY 648 LeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg 667  
 DB 8231 -----GGCGGCTGTACGACGAGGAGACCGAATACATGCACTG-----CCG 8190  
 QY 668 ThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThrTyr 687  
 DB 8189 AACGTCAAACTCACCGGCGCGCGAGTGGCGAGATCAACGAGCGGTGGAGTGCAGTCCG 8130  
 QY 688 SerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGlnAsp 707  
 DB 8129 TCGATC-----AACGCTTCGCGCGAGCGTGGCGCT--- 8100  
 QY 708 LeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu---AlaGlyVal 726  
 DB 8099 -----TTACGACAGTGGACAGCATTTCCAGCATCCGCTGTGACGACAGCTCCAG 8052  
 QY 727 AlaValHisLysThrValAlaAlaAspAsnLeuValAsn-----Promet 740  
 DB 8051 CTGATCTATCCGACGCTGCGAGCGCAACTACGACGTCGATGACCGCGCTGTCCCGCG 7992  
 QY 741 ArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAla 760  
 DB 7991 ACGGCGTGTCCGCGGCGAGTGTTCATCCGCGCGCGCGCGCGCGCGCGCGTGGACACC 7932  
 QY 761 AsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 780  
 DB 7931 AATTTCGGTCACTGACGCTACGCGCGACGCTAGCGTGTCTTGGCGCGCGCGCAATTCGG 7872  
 QY 781 GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
 DB 7871 GTGATCTGCGGTGGC-----GAAGGTGGCACCCACCTGACACAC 7833  
 QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysPheInsLeuArg 820  
 DB 7832 GATCTGTGGCGATGCGCGCGAGCTGCGGTCTTGGCGGTGGCCCAACAGCATCCGC 7773  
 QY 821 GlyTyrAlaHisAspSerLeuSerProIleSerAspLys-----GlyTyrLeuThrGly 838  
 DB 7772 GGCATGATTCGCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGGGT 7713  
 QY 839 GlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLys-----Asp 856  
 DB 7712 GCCAAGAACGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7656  
 QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyThrThrAsnAsp 876  
 DB 7655 TGGCGCGCGCGGTGCTGCTGCGATGCGGACGCTTGTGATGAC-----ACCCGCGAC 7602  
 QY 877 ThrLysIleGlyAlaGlyValGlyValArgTyrPheAlaSerProValGlyGlnValArgVal 896  
 DB 7601 TGGCGCAGCGCATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 7542  
 QY 897 AspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGly 916  
 DB 7541 GATATTGCGCGACGCTTGAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7482  
 RESULT 20  
 AEO12071/c 10839 bp DNA linear BCT 29-MAY-2002  
 LOCUS AEO12071 10839 bp DNA linear BCT 29-MAY-2002  
 DEFINITION Xanthomonas axonopodis pv. citri str. 306, section 449 of 469 of  
 the complete genome.  
 ACCESSION AEO12071 AEO08923  
 VERSION AEO12071.1 GI:21110627  
 KEYWORDS  
 ORGANISM Xanthomonas axonopodis pv. citri str. 306.  
 Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 Xanthomonas.  
 da Silva, A.C.R., Ferro, J.A., Rehnach, F.C., Farah, C.S., Furian, L.R.,  
 Quaggio, R.B., Montello, V., Montello, C.B., Van Sluys, M.A., Almeida  
 Jr., N.F., Alves, L.M.C., do Amaral, M.C., Bertolozzi, M.C.,  
 Camargo, L.E.A., Camarotte, G., Camarotte, F., Cardozo, J.,

TITLE	Comparison of the genomes of two Xanthomonas pathogens with differing host specificities Nature 417 (6887), 459-463 (2002)
JOURNAL	2 (bases 1 to 10839)
PUBLISHED	da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furian, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.P., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C., Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardoso, J., Chambergo, F., Clapina, L.P., Cicarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorfy, H., Faria, J.B., Ferreira, A.J.S., Ferreira, R.C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locall, E.C., Machado, M.A., Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spindola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezera, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Klajajima, J.P.
REFERENCE	Direct Submission
AUTHORS	Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900, Brazil
TITLE	Location/Qualifiers
JOURNAL	1. 10839
PUBLISHED	/organism="Xanthomonas axonopodis pv. citri str. 306"
AUTHORS	/strain="306"
	/db_xref="taxon:190486"
	/note="pathovar: citri"
	138. 665
	/gene="idnk"
	/note="XAC4197: gntv"
	138. 665
	/gene="idnk"
	/note="Identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
	/codon_start=1
	/transl_table=1
	/product="glucanokinase"
	/protein_id="AA039032.1"
	/db_xref="GI:21110628"
	/translation="MGVSGSGKTTIAQALAHYRFRFLDADDDYHVARAOMAGOPLTDAARLTPVVELSATLRDQVQAGSYVLASGLRSARHOLLRNSGVMRPVPIHAAPHVIARLSTRAGHEMPATLLDSQOTLEPLPDEADVSVVDASVPANVQAIQALAI DPAVQHTSRASEV"
	718. 1314
	/gene="XAC4198"
	718. 1314
	/gene="XAC4198"
	/note="putative; ORF located using Glimmer/Genemark"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="AA039033.1"
	/db_xref="GI:21110629"
	/translation="MRCALCVVACASRGYFRFADGQPAIARRAALPOCTEKTORRS AKPVALCASAHPRGMPKSTASATRHSLRLLPGGARGKRRVPPASRSPPHA VROATRSATSRSAEPISTPIRYIPQNAIDRRHRAPOSSCSPSRPSVAKAHTGCAATF
	1710. 3215
	/gene="XAC4199"
	1710. 3215
	/gene="XAC4199"
	/note="Identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
	/codon_start=1
	/transl_table=1
	/product="polyvinylalcohol dehydrogenase"
	/protein_id="AA039034.1"
	/db_xref="GI:21110630"
	/translation="MNSRPADEMKINRTTAKTLPAWTLTTGDISATPSVGGTLY VPMGGTLYSDVDTQGTGKPMWOKLSDYTGNAVSATVRPAISAKSIIVDDYAAGTLLA LDKKTGKLMKTIVANPOARITSSPVYGNRIYGVSSGDMGNLAHHTFESRGSV AVDLNMGKVMQFVYVPEVGTGASVGVAYIDPIINRRLYFGGNNYNIPOGVNCSN ARSYGSELTIVDELDCAHPDNYVDSVLSLNDTCKIYANAGCYDAMNLSQVAPF NGLCPNTQNTNLTGPDYDEGAGVOLFVPRNGRQOYLVAGOKSGITYWADAKGAK VMSVGPGGTGGVEMGTAPDPYKSOIYVAINNNSNATYTLGPNVATLHNGSMAAL DAATKTIKQYKVPBGINVPYVPAAGGSLIVSPNLLPAGSMAGNMVALADGATV WNTNAGSVYISSPAIANGALVWAGYGRFNFGTATGNQLTKEFVDNAAK"
	complement(3502. 3807)
	/gene="XAC4200"
	complement(3502. 3807)
	/gene="XAC4200"
	/note="putative; ORF located using Glimmer/Genemark"
	/codon_start=1
	/transl_table=1
	/product="conserved hypothetical protein"
	/protein_id="AA039035.1"
	/db_xref="GI:21110631"
	/translation="MATARRAICMGWRSMNMAARRSRAMSTGSPEROLHAYDE GYCGRPVLMGOLHSHVDSVSEGEDAALVALRRRAEAGLEAOIRLCQVTRL"
	complement(3710. 4180)
	/gene="arod"
	/note="XAC4201"
	complement(3710. 4180)
	/gene="arod"
	/note="Identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
	/codon_start=1
	/transl_table=1
	/product="3-dehydroquinate dehydratase"
	/protein_id="AA039036.1"
	/db_xref="GI:21110632"
	/translation="MLPEEVAMSIMIRGPEAATPIYRGPOPIPRAYIRQIVACAGDA GRIVALARCSGEOLDALRAAQAARWEIALIDGSCVDSARLHRYLRDLPYVETH DSDVDRPERCLPHGIGCIATVRYGCAOSITLGLTLELHNGCTEIOGDVHGT"
	complement(4170. 4550)
	/gene="XAC4202"
	complement(4170. 4550)
	/gene="XAC4202"
	/note="putative; ORF located using Glimmer/Genemark"
	/codon_start=1
	/transl_table=1
	/product="conserved hypothetical protein"
	/protein_id="AA039037.1"
	/db_xref="GI:21110633"
	/translation="MPMQLPACAGRALRTLACADVRLTAELHAAGDAVEVELYILDS GDPLSERSCARALRAAVDALPPYTELHTHNQDELPMLHQAHLAVLYITPHDPR AYAMSIGIARCLPSLHAPRVAA"
	complement(4859. 8419)
	/gene="XAC4203"
	complement(4859. 8419)
	/gene="XAC4203"
	/note="Identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
	/codon_start=1
	/transl_table=1
	/product="conserved hypothetical protein"
	/protein_id="AA039038.1"
	/db_xref="GI:21110634"
	/translation="MTLEPDLRPLLRKQLQLDAYGVSNATLNLGKSEEPFLTPSPNES

LPOINPLAIOADRIDVONLRITOLQPMIVLHKVOGELDVATGELTRDVIDTDMG  
DFRLHGDYLPNDYRADLDTAVLPAARGPASPGLAVAGDIDKMLAVLARGPAPL  
OASLVEFGDRDPWFAKAVPEALDTSLIPAAFGAONPAPPAITAINOASGGNAPL  
HGLKOGELSATLOPHIALADOVLVEPIVITFEERTOLKTRADRDONPSERRA  
VNASGRLFTPADLPATPDAPLVEVLDADLAVAGTILAAALIGRAVDEPQOQAEIV  
PDSRGNDRGAKOLQVOAKTPGSLIDLLEVAAPALOMDVAALAKRDEPFAAGN  
NLSGKIASKRQPLPAPAGVSPGFEEVLPSTLQROPALSGKFAVQOQGBEE  
LOLANSRINAKKYVDOLDIAEOLOPLDLDVLPATGVRLQOLVSKRDPADIT  
ADTAGLMDYTSANOISLRGLPMWGSQALNOSTLAEGVLDVSVQAOAGAE  
ALRDADIANMSAVALOGDVRNRPQOQVATIRIAPAKGMSIROAQBSTDA  
APTLSTCCATGATGALCASANPREGVAVHGDPLSLVQWPLPKOEGKOITLREL  
SLDGSFRPKNAMESRLASPEGJRLGETRYAAVANGNREGLIYDFOVSADPT  
OOQIOGKLFQFAGFVDAKFTYMDVAPLNGELTANSLRWLELVADIVYRPG  
LVGHVSLRGTDRKPLIGDADLTSDFAEYSPMGLTSEKREPRADDSAKITTA  
KSGPGLTVGSGMFGTSTPLLITNGDVAVNTESELITANPMQFOTDNTML  
RGLVYVPEADIDLERIDRGTSEVDYVLDVPEQCAPLMDLAIVIGDKYMSG  
FGLKGLSGOMQIRARPREMTANGCIDVGRKATQODLITRGOLITNNIYSDER  
VSLRAERKIDVTAGIDVSGRAESPRAVDSEPMSSGSEALSTYLGRGLSTASDET  
OQVSAASALASASSLIAQIGAKLGDGAVSQSSYLGVGVGKYLSPKLYGVYGV  
SMVGGSVLTKYLLSKRFDIEASSSTIETKGSYNWREK"

gene  
complement (8710..10491)  
/gene="XAC4204"

CDS  
complement (8710..10491)  
/gene="XAC4204"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/product="conserved hypothetical protein"  
/protein\_id="AA039039.1"  
/db\_xref="GI:21110635"  
/translation="MLKVAFLFLCTLEVPLOAVARATIDKIEIKLDDGDAEMIE  
NIOVSLIYAVGKEGESRLEYLALAEOTRELPFGYSPITILAAPRAGDKYT  
NIVTVRGEFVRVTHSITGMAEDORAYGODIKOEPREGGVESHPOVBAKVRIT  
RLAEGYRPADPTFORVATRAENADIDLMDSGRVMKRYRDXFRDLEPR  
LYWDGSTRYHEKRLRLSLKLDFTSTIDQKREEDQGVADVDKLAKRKT  
VYTSGLSYGSGAGVGVERRVYVNAHGKMDTQDLYAONRSLLTTSYVPAKRLMD  
GWYTSARLIDQETLEYIDLNRNKLGTGRSQIDNRMSALINLARMRWRSQDLR  
GAVETSTLIYPOLQANYVNVDDLFRPSVSAQMFTRGAEAGSDTNFNGOLYGLR

## Alignment Scores:

pred. No.:	7.75e-13	Length:	10839
Score:	359.50	Matches:	158
Percent Similarity:	35.54%	Conservative:	89
Best Local Similarity:	22.73%	Mismatches:	257
Query Match:	7.61%	Indels:	191
DB:	1	Gaps:	22

US-09-914-168-2 (1-919) x AE012071 (1-10839)

QY	235	ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle	254
DB	10272	CGGCAGACGCGT---GAGGCCCTGGAGCCGCTTGGCTATTCTGCGCCACCATTAACGCTG	10216
QY	255	IleArgAsnSerIleGlyGlu---ValAspValIleIleHis---AspLeuGlyGluPro	272
DB	10215	GCCGCACACACGCGCGGACAAAGTGACGCTGTGATCAGCGTGATCGTGGCGAGCGC	10156
QY	273	ValTyrIleAspTyrAlaGluValAlaArgGlyGluGlyAlaAspAspLysAlaIle	292
DB	10155	GTCGGGTGGCGACCTCGCATATCTCCATACCGGCTGGCGGACGACGCGCATCTG	10096
QY	293	ThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGlyLysTyr	312
DB	10095	GGCCAGATCTCAAGCAGTTTGGACCGCGGACGACGCGCATGCTTCAGCCATCCGACGTAC	10036
QY	313	GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisIleGlyTyrPheAspLysArg	332
DB	10035	GAAAGCCAGCAAGTCCCATCACCAGCGGTCTGGCGGAGCGCTGGTACTTGATGCGCAC	9976
QY	333	TyrLeuAspArgSerValAlaValIleLeuProAspAsnThrAlaAspValSerLeuIle	352
DB	9975	YTCACCCAGCGCGCTTGGATCAGCGGTGCGAGCATGCGCGCATCATGATCTGAAAT	9916

QY	353	TyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLys	372
DB	9915	TGGCACAGCGCGCGCGCTACACATGGCCAAAGTGCGCTTC-----	9874
QY	373	ThrAsnGlnLeuThrThrAspProAspLysLeuProValIleArgGluLeuLeuGln	392
DB	9873	-----GACTACAGACTACTTC-----CGGAGATGGCGCTTTCGACCGG	9838
QY	393	LeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsn	412
DB	9837	CTGGGTACTGGGAGGAGGCGAGCTACTACCAAGGCAACCTTGACCGCATGGCGGAG	9778
QY	413	AspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPheProGluArg	432
DB	9777	TGCGTGACCAACCTGGATTTACTTCCAGCACCATCATGATCCAA-----	9736
QY	433	GluGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThrGluPro	452
DB	9735	-----CCCAAGCCC-----	9727
QY	453	AlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGly	472
DB	9726	GAAAGCGCGAGCAGCAGCGCTGCTGCCGTTG---GACGTCAAACTCAC-----	9679
QY	473	IleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu	492
DB	9679	-----	9679
QY	493	AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAla	512
DB	9679	-----	9679
QY	513	IleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAspAlaValSer	532
DB	9679	-----	9679
QY	533	AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu	552
DB	9678	-----CGCGCC-----AAG	9670
QY	553	ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValPro	572
DB	9669	CGCACG-----GTCATCAACCTCAC-----	9652
QY	573	LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGly	592
DB	9651	-----GCTTGAGCTACGGC	9637
QY	593	SerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGly	612
DB	9636	AACGAAACGCGCGCGGTGCGCGCGGTGCGCGGAGCGCGGTACTGTAAACCGCGCGG	9577
QY	613	TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla	632
DB	9576	CACAAGATGACACACCATGATTAACCGGCAACCCGCAAGCGCTGACACCAATGAC	9517
QY	633	ThrLysProLeuSerHisAspProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln	652
DB	9516	CGCGTGGCGCGCTTCGTTGCTGGATGGCTGACACCGCGTCG-----GCG	9469
QY	653	GluValPheGlnHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGlu	672
DB	9468	CGGCTGTACAGCAGACACGACGAGTACATGACCTG-----GCGACGCTCAAACTCAC	9415
QY	673	IleSerArgSerIleIleGlnAsnGlyLysTyrAsnArgThrTyrSerLeuArgTyrArg	692
DB	9414	GGCACCGCGAGCGGCGAGATCAACAGAGCGCTGAGCGCATCGCTGCATC-----	9364
QY	693	LeuAspLysLeuLysThrGlnAlaProProGluGluThrGlnAspLeuProValAspPhe	712
DB	9363	-----AACGCTTGGCGCGCGCTTGGCGT-----TTC	9337

Db	Accession	Gene	Protein	Accession	Gene	Protein
Oy	713	ValAlaAspAsnLeuValAsn	-----PrometarglyTyraArgIn	745		
Oy	9336	AGCAGCGGCACGATTTCCGAAGCGCGGTGTACGAGACCTCCACACTGATATCCGCCA	9277			
Oy	732	ValAlaAspAsnLeuValAsn	-----PrometarglyTyraArgIn	745		
Db	9276	CTCCAGCGCCAACTACATCATGTCCGACGACCGCGCTGTTCCCGCGAGTGGCGTGGCGG	9217			
Oy	746	ArgTyrSerLeuValGlySerSerGlyLeuValSerAspAlasnmetalIaIaIa	765			
Db	9216	CACATGTTCTATTCCGCGCGCGCGCGAGGGTCCGCTCCGATACCACTTCGCCAGCTG	9157			
Oy	766	ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAlaTyrGlySerAsnArgIa	785			
Db	9156	TATGCCACCTGCGTGTGTTTGGCGGTGGGACAAATGGCGGTATCTCGCCCGG	9097			
Oy	786	HisGlnMetThrGlyAlaIleGlnAlaGlyTyrIleTyrSerAspAsnAsnHisVal	805			
Db	9096	-----GAGCGCGCGCACCCACTCGTGCAGCAGCAGCAGCTGTCGGCGAG	9058			
Oy	806	ProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAsp	825			
Db	9057	CGCCGAGCGCTGCGCTTCTTCGCGGTGGCGCCACACAGCATTCGCGGTATGCGTTCCGC	8998			
Oy	826	SerLeuSerProIleSerAspIlys-----GlyTyrLeuThrGlyGlyGlnValIleuAla	843			
Db	8997	GAACTCGCGCCACCGCACCGCCGACGATTCGCTGCGGTGCCAAGAACGCGTC	8938			
Oy	844	ValGlyThrAlaGlyTyrAsnTyrGlnPheMetIys-----AspLeuArgLeuAlaVal	861			
Db	8937	ACCGCAGTCCGATACACAGCAT---TACCTCAAGAGCGCGCCCTGGCGCGCGGGTG	8881			
Oy	862	PheGlyAspIleGlyAsnAlaTyrAspIysGlyPheThrAsnAspThrLysIleGlyAla	881			
Db	8880	TTCGTGATGAGGCGAGCGCCCTTCCATGAC-----CGCGCGACTGGCACACCGGTATC	8827			
Oy	882	GlyValGlyValArgTyrPheLeuSerProValGlyGlnValArgValAspValaIaThrGly	901			
Db	8826	GGCTTCGCGCTCGCTGCGCGCTGCGCGGTGGCGCGCGTGGATATTCGACATGCGC	8767			
Oy	902	ValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGly	916			
Db	8766	TTGACGACCGCCGACTCGCATTCACGCTCATCGATATCGCG	8722			
RESULT 21						
LOCUS	AY010120	20310 bp	DNA	linear	BCT 13-JUN-2002	
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						

REMARK	007, India
AUTHORS	Amino acid sequence update by submitter
TITLE	4 (bases 1 to 20310)
JOURNAL	Goel,A.K., Rajagopal,L., Nagesh,N. and Sontil,R.V.
	Direct Submission
	Submitted (15-MAR-2002) C.C.M.B., Habsighuda, Hyderabad, A.P. 500
	007, India
REMARK	Sequence update by submitter
COMMENT	On Aug 1, 2001 this sequence version replaced gi:11693112.
FEATURES	location/Qualifiers
SOURCE	1..20310
	organism="Xanthomonas oryzae pv. oryzae"
	/db_xref="taxon:64187"
	/note="pathovar: oryzae"
	218..1999
	/note="Orf1: similar to YTFM from Escherichia coli"
	/codon_start=1
	/evidence=not_experimental
	/transl_table=11
	/product="putative outer membrane protein"
	/protein_id="AAG38834.1"
	/db_xref="GI:11693113"
	/translation="MKVAFALFLLCTLPFLQAVNAATIDKIKELGDDDAEMETIE
	NIOVSISLRYAVGKEGCESERLFTLLAERKQTRALPEFGYSPITLAAADKADKVT
	VITLVDEGEFVRRTSTITGMWEDQRYLGDDLKQEPREGQVSHPOVEASKVRITIT
	RLRAEGRYFADPFTQRRVATIRAEHADIDILNMGSGRYDMGKRFEPDYFRDLFEP
	LVADESGYHSEEGKLDLRRESLTLDYESTIDIQPEAADQGRVVDVLTAKTRT
	VYTAGLSYSGESGAGVGRVRRYVNRGKMDQLDYANRKSITTSYRIPARMLD
	GWYTSARLYDEQTDYIDLNRVLTGSRSGINRMSAISIALRRMRSSDDFE
	GAVENTSLIYPOLQANYINVDRILFRPSGVSGMFTRGCAEGASPTNGOYIGQLR
	WELGAGNSRLILRGEGTITWTSDLVAMPSPSLRFAGGSINRQAVREGVPRPKPD
	KFLGAKNVYTSAREHYLVNGMGWGAIVVDSGASAPDRDWHITGIFGLRMSPVGE
	PVRVLGHNAGLNDPAQFOLYIDIGAN"
	2299..5853
	/note="Orf2: similar to yrfn from Escherichia coli"
	/codon_start=1
	/evidence=not_experimental
	/transl_table=11
	/product="putative periplasmic protein"
	/protein_id="AAG38835.1"
	/db_xref="GI:11693114"
	/translation="MKPDLRPLRLGRTLQNLNAVQVSNATNLGKSEEPFLPSMPESLP
	QLVPLAIDAKDIDVONLRITTOQPMPIVAKVOGGLDVTGALRTGDIYIDMGEP
	RLGMDVPSPDDYHADLTATVLPAAERRPASIGIARGLDMKEVIAIARAPALQIA
	SIVESGRDDPTMAFKAVTEALDPTSLITPABEQANPAPITIALNLQASGSGNADLG
	ITKQAFSATLPQSHIALADQVLTPEPLVIDTEEGRQLKGTADLRPTQNSPFEAVN
	ASLRTLPADPAPTPADPLVPEVLKDAKRLGAVAGTLAKAMAIGATVERDQOQLVED
	SRKNDRAQLQKQVATPGGTLDLSGVNAWPELQMVNQLKDPGYAPMGNGL
	SGIASKGGQLPAPAGGVSPGFETAPEVSLTQGLRQALSAHGKFKALRQEGSGEIL
	LALGNRIAKKGLDQOLDIAAQLQPLQDLDDVPLAGTVGVRQLOVSGKRDADPDITD
	IANGRLMPTYSNOSTSLGRRLPMRSGDQLAQGRATIEGVYLDVSRVQARAAVEL
	RLIADIANMSASVALOGDVRRNCRNOGVATIRIAPSKDDANAALRQPAFSTDGAF
	TLSDTDLGATGALCALPAAVPRAGVNVGHDAPLSLVQWPLRQKQGRQLYLRGELS
	DDSEFKRGANAMEGSLIASAPBEGIRLGETRYAAVAGNPNGELRLDOESMQADFTQ
	QILOGKIGIGFOGAGFVDAKFNTEMDVAPLNGELIYINMSLYWLEIYIADVPRGFR
	EGHVSJLRGTDRKPLGSDATILSDFTAEYSPMGALIEBGRKFPALPDGSAKITASKS
	GGTLLVDGGLSMFETSTPLINRGDNYAVTSLRTLIANDMOFGITDNTMHLRG
	KTVPEADIDLERLDRGCTSYSEVDVYLDVDPDQRTASPLDMDLAVLGVKVMGTG
	LKGLSGQMKIRARPEREMTANGLDVDRGKRYKRGODLTITRQLTMNHNIVSDPRVS
	LRAERKIGDITAGIDVSGRAESPRADVWEPASQSEALVYLGVLGSLVAASDEQTO
	VASASALAGSGSLIASQIGATGLDEACVOSGTSGLSVYFGFKYISPKLYGVGYVSM
	ICGGSVLTLYKLISRCFDLEAESVETVEKGSVMWREK"
	106344(5988..6934)
	/note="tnpxo: nonfunctional putative transposase due to
	frameshift mutation"
	complement(7188..7625)
	/note="Orf3"
	/codon_start=1
	/evidence=not_experimental
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAG38836.2"
	/db_xref="GI:19483794"



CDS  
/translation="MGRDAIATLIPHQSMCIWEDVIAQRIVLRSHNRSDAP  
LRSNGRLALHLCXEYGAQAMAVGGLGASGRKPRPGMLVALRGVELHVTLETPE  
AIDCEAOVLMQESDSQYSEFRLHGEQLLEGRATVMLGAQA"  
complement(7612..10035)  
/function="probably involved in transport of xanthomonadin  
to the outer membrane"  
/note="Orf4"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="putative inner membrane protein"  
/protein\_id="AG38837.1"  
/db\_xref="GI:11693116"  
/translation="MPRLRPACVRPLQODTHWALELNARRIGLALLLAVAGF  
MLSELTAVTGDRKMPAPRTPAKOGLIELEGPGSRLLMALNSDPYTLAEOSQ  
LOKALAAOPELFEIVNGNGNAGIEAIKRLPRYLRLSDSPDKTPLDPALEAOL  
VODGSRAPAAWYERLLPRDPTMETILHLETLPAAAPQTRNDYFPRACISALLIYQ  
RAAGFDPTGOQLAYDAVQAFAKVSKESTRLTLTPGAFVAEITARTQEAQWIGTL  
DTVGLVLLLVAVRSWKIPVLGVLPLASAGLAVAVLFDVHGIVAFGFTLIGV  
VODYPILHFSHQPGIDPRENARHMLPTLVTVGVSTCIAYTFLEFGVDGLAVET  
IGLATAAATTRMLPALIDPAPADYADSRLLAVMRGVARLPRPRLVLSAAVAVSI  
AVILAPGOFMONDSKLTPVPEPKALADOMHILROELGADPRVYALAPANDAAVIA  
SROLRPRDLAVRNQDLGIDMAARILPSAKTOHARQALPDAQAQATETEVATTP  
FNSDAFAPFLQDLDAKKNAPLPLKDTGSPLATSVGGLLGGDRSTALVSLTGLRD  
PAVLAAYVGGSGAQLDLKDASELSIAAARGVLAALAVLAALAVTVALSLSPRI  
VVLPLMALTVLILALIRGTGVELMLFHLILALAGLGLDALFEDHAGDHDADOL  
RTHALIVCSLMTLILVFALLAASSIPVLRAIGSTVALGVLFNFILALVSRPALERT  
ANPEHPAGP"  
complement(10120..10767)  
/note="Orf5"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AG38838.1"  
/db\_xref="GI:11693117"  
/translation="MPMCRPLPMLLILFSAAPLCAAPPETLDGVLELRAPAVS  
TEFVELRGSSLLKTPLRVGHYRRPDQOLVREVSYPGHEHTTLTGDEGLEREKSP  
RRFSLSRVPELAGLKGSGFGLALSGDRALLOHNSVSGGPOPMALLTDQKDAVVKO  
VRQLRLYGRKDELRCIESVPARKGVORTLLAAGARAAAGVTATATLTLCHSGSA"  
complement(10691..11665)  
/note="Orf6"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="putative acyl transferase"  
/protein\_id="AG38839.1"  
/db\_xref="GI:11693118"  
/translation="MSKHWKQREGGGRPALMLIRGIAIHGRGVNRRALYPTLYFL  
LVRAPERASRAVYLARVILGRPATLRDVAARHITFASTILDRVYMLGCOMQRFVDSIG  
LDQLHTQMDRGSGVLIFGSHLGFALRVLATERPVPQKVVYDKAHNRMTDLGLAL  
NFOLANIIDACMDATSIYMAIKRATDAGALVALLLDRPREGPARMLAVLMDVAA  
TSPMLIAALKVYVLAIFGLYHGNHVELAFEFSSGLDPRHQBARMLAVLMDVAA  
RLEHTCYAPYWMFNYPFVNTHHADVPHPAADATVAVQRTAMRRTA"  
complement(11662..11949)  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="putative dehydratase"  
/protein\_id="AG38840.1"  
/db\_xref="GI:11693119"  
/translation="MDFVYHDPCLPGHFGKRPVYGVVLDHVLQAVEALHGPRAA  
ARLPQVKFVQPLPGQTSVTLLEGDGRPMRFRVQGRADVVLVSGELVAEAGT"  
complement(11961..12710)  
/note="Orf8"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="putative acetyl transferase"  
/protein\_id="AG38841.1"  
/db\_xref="GI:11693120"  
/translation="MPLARPSRLAAVNNLQLAFTLAFAGIYVALALMAITPDA  
DMLRMSGWMAAPVLEFHGAGVTYIVGERERVMSKNYLFVSNHQSIIIDICALFYALPV

CDS  
PLRFLKEEMLKVPFVNMVATGMLFLDRDHRBAGALVROVAALLROGKQLCPPE  
GTRSRGTALLPFEKYLIDAGVOVVPVALDCCGVLPVDFLFRVPGIIVRICT  
PIAVTGGDAPDRQOLTEQAKKAVAMHAPRI"  
complement(12764..13438)  
/note="Orf9"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AG38842.2"  
/db\_xref="GI:19483795"  
/translation="MGVLLAVASPLAHMANASHRPEVAYIAGLLVLMVLEPMVAR  
RAMAMALALLLAGLVLMHSPYAMVLAAPVYTGVMVAFEGSRSLRHGTPLITFI  
VEGLYGGAGMVPTEORRTYRLTLAMALLLCGLTLNVLVGLCAEESGYTOGVAS  
PLPSDARSLFANLLCYIGGFEIEYLLRGWMPQRPYRNLPDFLRQMARLGRPF  
WRDLR"  
complement(13485..13883)  
/note="Orf10"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="putative acyl carrier protein"  
/protein\_id="AG38843.1"  
/db\_xref="GI:11693122"  
/translation="MLGLAVATATTTADAPRTSDGRPHCCCLFLILPPYDLKSMW  
SQTAAERELAEVLVESLLEEDVOPADIDEAPLPLNGLGIDSLDALELALISKRYG  
FQLRSDNDENRIRIFASLRALSAHNAKTY"  
CDS  
Alignment Scores:  
Pred. No.: 4,24e-12 Length: 20310  
Score: 353.50 Matches: 153  
Percent Similarity: 35.11% Conservative: 91  
Best local Similarity: 22.01% Mismatches: 260  
Query Match: 7.48% Indels: 191  
DB: 1 Gaps: 20  
US-09-914-168-2 (1-919) x AV010120 (1-20310)  
QY 235 ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle 254  
|||||  
DB 437 CGGCAGACGCCG---GAGCGCTGGAGCCGTTGGCTTTACTGCCCAACCATCAGCTG 493  
QY 255 IleArgAsnSerIleGlyGlu---ValAspValIleIleHis--AspLeuGlyGluPro 272  
|||||  
DB 494 GCTGGCCGACCGCGCGGACAGAGTGCACCGGATGACAGCGTGCATCGTACAGCGCG 553  
QY 273 ValYrIleAspTyrArgAlaValAlaGluValArgGlyGluGlyAlaAspAspLysAlaIhe 292  
|||  
DB 554 GTGGCGGTGGCCACCTGCATATTTGCATACCGCGCTGGCGGACGACGACCGTTATCTA 613  
QY 293 ThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGlyLysTyr 312  
|||  
DB 614 GCGCAGACCTCAAGCACTTGCAGCGCGGCGGACGCGAGCGCTTGTAGTCATCCGACATAC 673  
QY 313 GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluIleGlyLysTyrPheAspLysArg 332  
|||  
DB 674 GAACCGACGACGATGCGCATACCGCGCGCTGGCGGACGCGTACTTGCATCGACCGAC 733  
QY 333 TrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIle 352  
|||  
DB 734 TTTCCCAACCCCGCGATGCGATCAGCGCGCGGCGCGCGCGCATATCGATCTGAAT 793  
QY 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLys 372  
|||  
DB 794 TGGACACGCGCGCGCTAGACATGCGGCAAGGCGCGCTC----- 835  
QY 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGln 392  
|||  
DB 836 -----GACTACGACTACTTC-----CGCATGCTCTTGTTCGACCGC 871  
QY 393 LeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsn 412  
|||  
DB 872 CTGCTGACTGGGACGAGGCGAGCTACTACACAGCAAGCGAGCTGGATCGGCTACGCGAG 931

QY 413 AspLeu1leAlaThrArgTyrPheAsnMetValAsnThrGluLeuValPheProGluArg 432  
 Db 932 TCCTGACCAAGCTGGATTACTTACACACATC----- 964  
 QY 433 GluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGluPro 452  
 Db 964 ----- 964  
 QY 453 AlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGly 472  
 Db 964 ----- 964  
 QY 473 IleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492  
 Db 965 -----GATATCCAGCC----- 976  
 QY 493 AsnLeuValAlaAlaLysAlaThrArgHisLeuTyrAspMetProAspAspArgValLeuAla 512  
 Db 976 ----- 976  
 QY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGluArgIleSerAspAlaValSer 532  
 Db 976 ----- 976  
 QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552  
 Db 977 -----AAGCCGGAAGAACCGACGACG----- 1000  
 QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLys---LysVal 571  
 Db 1001 -----GGCCGCGTGCCTGGAGCTCAAACTACACCCGCGCAAGCCG 1042  
 QY 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyr 591  
 Db 1043 ACGCTCTACAC----- 1069  
 QY 592 GlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAsp 611  
 Db 1070 GCGCAGCCAGCCGCGCGGTGGCGCGGTGGCGCGCTTACGTAATGGCGCT 1129  
 QY 612 GlyTyrGlnAlaGlyAlaGluLeuArgLysSerGluAspLysGlyLeuValLysLeuTyr 631  
 Db 1130 GCGCCACAGATGACACGACGCTGATTCAGCCAGAACCGCAAGACCTTGACACACAC 1189  
 QY 632 AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGln 651  
 Db 1190 TACCGCATACCGCGCTTCGCTGGCGATGGCTGACACCGCTTCG----- 1237  
 QY 652 GlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671  
 Db 1238 GCGCGCTTACGACGACGACGACGATTAATCAATCACTG-----CGCAACGCTCAAACTC 1291  
 QY 672 GluIleSerArgSerIleIleGlnAsnGlyLysTyrPasnArgThrTyrSerLeuArgTyr 691  
 Db 1292 ACCCGCAGCGCGACGCGCATCAACAGCGCTGACGCGGATCGCTGCATC----- 1345  
 QY 692 ArgLeuAspLysLeuLysThrGlnAlaProProGluThrTyrPgnAspLeuProValAsp 711  
 Db 1346 -----AACGCTTGGCGGACGCTTGGCT----- 1369  
 QY 712 PheValAsnGlyLysProSerGlnGluAlaLeuLeu---AlaGlyValAlaValHisLys 730  
 Db 1370 TTCAGCACCGCGACGATTCGAGCGCGGTGTACAGACCTCCACGTTGATCTATCCG 1429  
 QY 731 ThrValAlaAspAsnLeuValAsn-----PrometArgLysTyrArg 744  
 Db 1430 CAGCTGCAGCGCACTACATCAATGTGCAGACCGGCTGTTCCCGCGCAGCGCGGTGC 1489  
 QY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764  
 Db 1490 GCGCAGATGTTTCATCGTGGCGGTGCGCAGCGTGGCGGCTCCGATACCAACTTGGCCAG 1549

QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
 Db 1550 GTGATATGCCAGCGTCCGTTCTCGCGCGTGGCAACAGCGGTTGATCTTGGCC 1609  
 QY 785 AlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleThrPheAspAsnHis 804  
 Db 1610 GGC-----GAAGCGGCACCACTGACGACGACCTGTGCGC 1648  
 QY 805 ValProTyrArgLeuArgPhePheAlaGlyIleAspGlnSerIleArgLysTyrAlaHis 824  
 Db 1649 ATGCGCCCAAGCCTGCGCTTTCGCGGTTCGCAACAGCATTCGCGTTATGCTTT 1708  
 QY 825 AspSerLeuSerProIleSerAspLys-----GlyTyrLeuThrGlyLysGlnValLeu 842  
 Db 1709 CGCGAAGTCCGCGCGCGACCGCCAGCCGACACTTGCCTGGGTGCGCAAGACGTG 1768  
 QY 843 AlaValGlyThrAlaLysTyr---AsnTyrGluPheMetLysAspLeuArgLeuAlaVal 861  
 Db 1769 GTTACCGCCAGTCCGATTCGATCAACAGCGTGGCGCGTGGCGCGCGCGGTG 1828  
 QY 862 PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAla 881  
 Db 1829 TTCGTGACAGCGCGCGCGCTTCGACAC-----CCCGGACTGGCACACCGGTATC 1882  
 QY 882 GlyValGlyValArgTyrAlaSerProValGlyGlnValArgValAspValAlaThrGly 901  
 Db 1883 GCGTTGCGCTCGCTGTTGGCGTGGCGCGGTGGCGCGGTGATATCGCATATCGC 1942  
 QY 902 ValLysGluGluGlyAspProIleLysLeuHisPhePheIleGly 916  
 Db 1943 TTGAACGACCCGACGCGCGCATTTACGCTTACATCATTCGCG 1987  
 RESULT 22  
 AEO03957  
 LOCUS  
 DEFINITION  
 AEO03957 Xylella fastidiosa 9a5c, section 103 of 229 of the complete genome.  
 VERSION  
 AEO03957.1 GI:9106207  
 KEYWORDS  
 Xylella fastidiosa 9a5c.  
 Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 Xylella.  
 SOURCE  
 ORGANISM  
 Xylella fastidiosa 9a5c.  
 Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 Xylella.  
 REFERENCE  
 1 (bases 1 to 10074)  
 SIMPSON, A.J., REINHARD, F.C., ARRUDA, P., ABREU, F.A., ACENCIO, M.,  
 ALVARENGA, R., ALVES, L.M., ARAYA, J.E., BALA, G.S., BAPTISTA, C.S.,  
 BARROS, M.H., BONACCORSI, E.D., BORDIN, S., BOVE, J.M., BRIONES, M.R.,  
 BUENO, M.R., CAMARGO, A.A., CAMARGO, L.E.A., CARRARO, D.M.,  
 COLAUNHO, N.B., COLOMBO, C., COSTA, F.F., COSTA, M.C., COSTA-NETO, C.M.,  
 FACINCANI, A.P., FERREIRA, A.J., FERREIRA, V.C., FERRO, J.A.,  
 FRAGA, J.S., FRANCA, S.C., FRANCO, M.C., FROME, M., FURLAN, L.R.,  
 GARNIER, M., GOLDMAN, G.H., GOLDMAN, M.H., GOMES, S.L., GRUBER, A.,  
 HO, P.L., HOHEISEL, J.D., JUNQUEIRA, M.L., KEMPER, E.L., KITAJIMA, J.P.  
 AND MARINO, C.L.  
 The genome sequence of the plant pathogen *Xylella fastidiosa*. The  
 Xylella fastidiosa Consortium of the Organization for Nucleotide  
 Sequencing and Analysis  
 Nature 406 (6792), 151-157 (2000)  
 MEDLINE  
 20365717  
 JOURNAL  
 10910347  
 PUBLISHED  
 REFERENCE  
 AUTHORS  
 2 (bases 1 to 10074)  
 SIMPSON, A.J.G., REINHARD, F.C., ARRUDA, P., ABREU, F.A., ACENCIO, M.,  
 ALVARENGA, R., ALVES, L.M.C., ARAYA, J.E., BALA, G.S., BAPTISTA, C.S.,  
 BARROS, M.H., BONACCORSI, E.D., BORDIN, S., BOVE, J.M., BRIONES, M.R.S.,  
 BUENO, M.R.P., CAMARGO, A.A., CAMARGO, L.E.A., CARRARO, D.M.,  
 CARRER, H., COLAUNHO, N.B., COLOMBO, C., COSTA, F.F., COSTA, M.C.,  
 COSTA-NETO, C.M., COLAUNHO, L.L., CRISTOFANI, M., DIAS-NETO, G.,  
 DOCEÑA, C., EL-DORRY, H., FACINCANI, A.P., FERREIRA, A.J.S.,  
 FERREIRA, V.C.A., FERRO, J.A., FRAGA, J.S., FRANCA, S.C., FRANCO, M.C.,  
 FROME, M., FURLAN, L.R., GARNIER, M., GOLDMAN, G.H., GOLDMAN, M.H.S.,  
 GOMES, S.L., GRUBER, A., HO, P.L., HOHEISEL, J.D., JUNQUEIRA, M.L.,  
 KEMPER, E.L., KITAJIMA, J.P., KRIEGER, J.E., KUTAME, E.E., LAIGRET, F.,

Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A.,  
Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N.,  
Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L.,  
Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C.,  
Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A.,  
Nascimento,A.L.T.O., Netto,L.E.S., Nham Jr.,A., Nobrega,F.G.,  
Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C.,  
Palmeri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira  
Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V.,  
de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V.,  
Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva  
Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de  
Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M.,  
Tsunako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S.,  
Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.  
Direct Submission  
Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and  
Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP  
13083-970, Brazil  
FEATURES  
source  
location/Qualifiers  
1..10074  
/organism="Xylella fastidiosa 9a5c"  
/db\_xref="taxon:160492"  
/clone="9a5c"  
complement(271..759)  
/gene="XF1230"  
complement(271..759)  
/gene="XF1230"  
/note="similar to SPIP76270 (percent identity: 49 %/query  
alignment coverage: 87.7 %/subject alignment coverage:  
77.6 %); identified by sequence similarity; putative; ORF  
located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA84040.1"  
/db\_xref="GI:9106208"  
/translation="MPLAPPIITGNKPEYQALQEAQAKALIHGETDRIANAALIAL  
IFHSLPGLMWGEFYFDGRELWVGFQGLPACIRIPIDKVGCTAARTRRQRIIDVD  
TFPDHICDSASRSELYVPLKDDTLIGVLDLSDPTCARFSDQGTGIESLARFVSS  
LR"  
1052..2905  
/gene="XF1231"  
1052..2905  
/gene="XF1231"  
/note="similar to SPIP39320 (percent identity: 27 %/query  
alignment coverage: 95.9 %/subject alignment coverage:  
102.6 %); identified by sequence similarity; putative; ORF  
located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AA84041.1"  
/db\_xref="GI:9106209"  
/translation="MKSRCFWNVRFITLSMLPRYAQMLRRIIPMLISMLFASLASLA  
QAVGVKIQIKGDKGDVEMKNNIVSLSYVSGKEGEFLLEYLLSAGEQTRQ  
ALPEFGYAPTRIDARONDHITIVYVDGEPVRAQAVAMTGAADHAYDIL  
EDRPKIGELIFNHPYASKVIRITRLAERGVFPADFTHRVEVTRAHADHIDILIME  
SGRYDMGVPVRFHYDFHEGLFNPLVYWDGSSVGECKIDRLRESLTKDPESSDIO  
PKREBAPBEGNVPYDVKLERKSKITYAGISYGSAGCLAGACVBRMRNARGHKMA  
RLDYAOMLKSUTRAYQIPAFKWLDCWYIFARAYDEQTKYIDLNRNKLISARSGQINR  
HLTATASLNLKGRMRYRADGNTVAYQOSTLVYPOLEASVYDVDAIFPNSAAT  
VILRGASALGSKSNFTQLGOLRFGHGLGASRLILIEAGTWTASDLVAMPISAR  
FAGGVANSIRGYAFREYVGPRTAKPDPAFALGAHVSAGAEFEYFKGGPFGGAVFDSC  
SAFNRYPDHMTGIGILRYSRSPGVAVDLARGLNSDSKIQTLIDIGANL"  
3184..6723  
/gene="XF1232"  
3184..6723  
/gene="XF1232"  
/note="hypothetical protein; identified by sequence  
similarity; putative; ORF located using  
Glimmer/RBSfinder/Start codon shift: 411"  
/codon\_start=1

/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA84042.1"  
/db\_xref="GI:9106210"  
/translation="MDPDWLLLRLLRDVLEMSNASISLCOSEEPNFPSPDPLP  
SINPLAQLDITITIDNFRISLKQPLHLRIKGAVANVADDELKIQRLAVSDRONI  
RADGVYRRQDSTDLTATVAPLSVRGPASGFTVIRQDLDELQIALAGRAPAPQA  
DVLVGRSAPRMTFSAASEHFFPGLPELRTPTSATLALNMOACGGGADLQQA  
FKQGGOKIVLOSRLQALNOMVMPVIVDAGRTNLRNGKVDLPARTRHFPSNA  
RGLSPKSPBEPNTQVRAVVPVLEAGVSGSTLIRANMAAGARLTRDAQRLIFDY  
RGDAHATFKQVQATPGCLDFTGCVNAPALRMDLAKELQDFGTFIPQMWNV  
GOLTSQGRQRYVSSGIEAIVAPGLNGELRQAPMOASGFTFQDQGEAFQIGDS  
HISAKGVPKPHLDLAVQOSLNLNDLPSKEGSLILRLRPAHPVTVLNGTSL  
RNMADRQKQFALHGLPMRNGKLTLOAEIEGTLDSVSIHAGAVENLQADV  
RHAMLLHLHGVRHGDARWREDLSQLRREPDKGAMALRAPOVYMANPVFTLSATC  
LDAOGGALCLHADPHGICLVKGOALPLTIVYPMPLPKNGRSINLRGALTLDGHRP  
HDAWEGSFRRASSLEGTRGCPRIADVAGVDRELYNRDQSVQVDLTSGHIVKL  
GIFQNGFVDAKFMVMDPPAPLPGDLYMNSRLYWMEMPDIQNPGLLEGHVL  
RGTRSGPLIGEATLRDFTAEYPSMGLNLSOGGRTAQPDGSARILASVRGDTLN  
IDGGLSMFGENNPIQLNIRNVAVNTSELRIVSDPLRFGLINNTMQLQGEAVSE  
ADIDLERLDRGYSVEDVYLDPPVPERTAVSPLEMNLRIVGDVYNMSPGLKGLS  
GCOMIARPPGHMTANGGLDVGRRKAYAGODLTIRGOLVYNNITVSPRVATRRQR  
IADITRAGIDVSGRAQSPRYQWSDPVMSQSELSLVLRKLTGTSASDENOQVASAA  
ALSAGGLIASOLGAKLGLDAGISQSSITLGGSVGFKTYLSPKLYGVGVSMVSGS  
VLTLYLLRSGFDLEVSSTVETKGSVMRKER"  
complement(6826..6945)  
/gene="XF1233"  
complement(6826..6945)  
/gene="XF1233"  
/note="hypothetical protein; identified by sequence  
similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA84043.1"  
/db\_xref="GI:9106211"  
/translation="WLFPEHHIGNSAMSLAIGASFLVIDPQFTGQANRCP"  
7500..8396  
/gene="XF1234"  
7500..8396  
/gene="XF1234"  
/note="similar to SPIP77541 (percent identity: 62 %/query  
alignment coverage: 98.3 %/subject alignment coverage:  
99.0 %); identified by sequence similarity; putative; ORF  
located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=11  
/product="carboxyphosphonoenolpyruvate phosphonmutase"  
/protein\_id="AA84044.1"  
/db\_xref="GI:9106212"  
/translation="MSTSLPIQTPGQCFPTALMTERPLOVICAMNAYHALLAORAGYR  
ALYSGGVAAGSLGILPDGMNTLEDVILDIRINDVCDPLVINDPGEFGSANIIE  
RIKTLIRKAGAACHITENOVAKKRGCHRGKATVLSIDMADRYKAAADKRTAAFLFI  
AKTDVIAQEGVSAIERAOAYVEAGADSTIFAAVSDLETYRHFVERIRVPLANNTEF  
GKPTPLTREOLEAVAGVAIOLPFLSAFRANKAAETVYNAIRRDHQOAVLPIMQTRAE  
LYERDIAHAFEGRLDALASRS"  
complement(8703..8778)  
/gene="crNA-Lys-1"  
complement(8703..8778)  
/gene="crNA-Lys-1"  
/product="crNA-Lys-1"  
/note="Found by trnscan"  
9291..9446  
/gene="XF1235"  
9291..9446  
/gene="XF1235"  
/note="hypothetical protein; identified by sequence  
similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA84045.1"  
/db\_xref="GI:9106213"



OY		685	ArghrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProFroGluThr	704
Db		285	TATTTGGCGGACTTCGAGCAAGAATACATAGTTTT---ACACAAGCGCAT-----	332
OY		705	TrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuLeuAla	724
Db		332	-----	332
OY		725	GlyValAlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyTyrArg	744
Db		333	-----ATCACAGATAAAACCCTTA-----CTTCCTTTATCCAACAGTGTTGGATTTCCT	377
OY		745	-----	749
Db		378	CGTACTCGATTACGGTGTGGTTCCTTTGCCACTTGSGGCGATGTCCAAAAATTTACTTTT	437
OY		750	GluValIGlySerSerClyLeuValSerAspIleAsnMetAlaIleAlaArgAlaGlyIle	769
-Db		438	GATTTAAGCAACAGCAATTTGGCTATACAGATCTTTTATTAAGAAGCAGACCAACAGC	497
OY		770	SerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThr	789
Db		498	GCCTGGATTCGACTATTGACGAANAAT-----CATCGATTGCT	536
OY		790	GlyGlyIleGlnAlaGlyTyrTlleTrpSerAspAsnPheAsnHisValProTyrArgLeu	809
Db		537	GCTCGTCGTGAAMTCGGGTATTACATACAAAAGATATTGAAAATAATCCGCGCTACACTG	596
OY		810	ArgPhePheAlaGlyAspGlnSerIleLeuArgGlyTyrAlaHisAspSerLeuSerPro	829
Db		597	CGTTCTTGTCGGGGGCGATCCAGTGGCTGGTTACGGCTATTAATAAAATTCGCGCT	656
OY		830	IleSerAspLysGlyTyrLeuThrAlcylGlyGlnValIleuAlaValAlcylThrAlaGlyTyr	849
Db		657	AAAAATAAAATGGAANAATTTGGTGGCTGACGTTTGCTTAGACGTTCTTTACAAATAT	716
OY		850	AsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspTlleGlyAsnAlaTyr	869
Db		717	CAATATCAAGTTTATCCGAATGCGTGGCGCACACTTTTTCGAAATAGTGGATTACCCGCT	776
OY		870	AspLysGlyPheThrAspAspThrLysIleGlyAlaGlyValAlcylValArgTrpAlaSer	889
Db		777	AATAT-TCACACGCAAAAGAGCTGCGTTATGGCGACAGGCGTTGGTGGTGGCGATCG	835
OY		890	ProValIGlyGlnValArgValAspValAlaThrGlyValLysGluGlyLysAsnProIle	909
Db		836	CCAGTGGCGCGCATTAATTTGATGTATGTGCACACCCATTGCGGATTAAGAATATACGCAA	895
OY		910	LysLeuHisPhePheTlleGly	916
Db		896	AATATTCAATTTTACATTGGA	916
<hr/>				
RESULT 24				
LOCUS	AL646064	201050 bp	DNA	linear BCT 07-DEC-2001
DEFINITION	Ralstonia solanacearum GM11000 chromosome, complete sequence;			
ACCESSION	AL646064	AL646052		
VERSION	AL646064.1	GI:17428340		
KEYWORDS	.			
SOURCE	Ralstonia solanacearum.			
ORGANISM	Ralstonia solanacearum			
REFERENCE	1 (bases 1 to 201050)			
AUTHORS	Salanoubat,M., Genin,S., Artiguenave,F., Guzy,J., Mangenot,S., Arlat,M., Billault,A., Brotilier,P., Camus,J.C., Cattolico,L., Chandler,M., Choise,N., Claudel-Renaud,C., Cunneac,S., Demange,N., Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schlex,T., Signier,P., Thebault,P., Whalen,M., Winker,P., Levy,M., Weissenbach,J. and Boucher,C.A.			
TITLE	Genome sequence of the plant pathogen Ralstonia solanacearum			

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 201050)  
AUTHORS Boucher,C.A.  
TITLE Direct Submission  
JOURNAL Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston  
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie  
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,  
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean  
Dausset-CEPH, 27 rue Juliette Dedu, 75010 Paris, France, LMG CNRS  
118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA  
URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,  
Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,  
F3136 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire  
INRA, BP27, F31326 Castanet-Tolosan Cedex  
Christian.Boucher@toulouse.inra.fr  
http://sequence.toulouse.inra.fr/R.solanacearum.html.

FEATURES  
source Location/Qualifiers  
1. 201050  
/organism="Ralstonia solanacearum"  
/strain="GM11000"  
/db\_xref="taxon:305"  
206. 1447  
/gene="RSC1326"  
/note="RS02848"  
206. 1447  
/gene="RSC1326"  
/EC\_number="2.6.1.1"  
/function="small molecule metabolism; amino acid  
biosynthesis"  
/note="Product confidence : probable  
Gene name confidence : hypothetical  
Predicted by Codon\_usage  
Predicted by Homology  
Predicted by Framed"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="PROBABLE ASPARTATE AMINOTRANSFERASE PROTEIN"  
/protein\_id="CAD15028.1"  
/db\_xref="GI:17428341"  
/translation="MKTIQSAKLNVCYDIRPVLEKAKOMEESGHOITKLNIGNLA  
PGFPAPEIIOODMIRNLPNSAGYSDSGIPAPAKAMVHYTOOGIKNTLIDYIGN  
GASLLIALATNALDLADGDELLPAPYPLMTAMTSLSGTPVYILCDENSGMRPDD  
IRAKTPTKGTIVINPNPRTGALISDALHIDIAIAREHGLVIFADEYDKVLPDND  
RHTAMASUEDVLTFTFNSLSYNSCGIRAGMNVSGDKRPADKDIETGLNLSMRL  
CANVCGQWAIQTALGYSINDLVAPGMRQRDLAYELIAIPGVSCVKPAALYM  
FPRDPSTVYIDDQTFIRQLLEERVLVVOGTGFNMHSPDHFRIVFLPHEDDLREAI  
GRIARFLERYRQRHGTGIRAA"  
1538. 2857  
/gene="RSC1327"  
/EC\_number="1.1.1.3"  
/function="small molecule metabolism; amino acid  
biosynthesis; threonine"  
/note="Product confidence : probable  
Gene name confidence : hypothetical  
Predicted by Codon\_usage  
Predicted by Homology  
Predicted by Framed"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="PROBABLE HOMOSERINE DEHYDROGENASE OXIDOREDUCTASE  
PROTEIN"  
/protein\_id="CAD15029.1"  
/db\_xref="GI:17428342"  
/translation="MKPIKIGLLGCTVGGGLTKVLQNRQEIIRRRAGGIEVALIAV  
RNVERRAIVQETGASVPVDDPAVAVDSPEIDIVVELIGVDLTRELVLRAIANKH  
VYTKAKALLAEGNEIFKAADKGVAVAFEAIVAGIGIPIIKALREGLTNRIMELIGI  
INGTNFLTSEMRKDGCTGVLAQORIGVGAEDPTDIEGFDAAHKLTLMISALAFG  
VPVQERAVESITKLEAVNDIRYAEELGRIKLLDLTRAEAGIELRHVPTLVPRRL

gene  
CDS  
2889. 4334  
/gene="thrc"  
/note="RSC1328; RS02850"  
2889. 4334  
/gene="thrc"  
/EC\_number="4.2.99.2"  
/function="small molecule metabolism; amino acid  
biosynthesis; threonine"  
/note="Product confidence : probable  
Gene name confidence : probable  
Predicted by Codon\_usage  
Predicted by Homology  
Predicted by Framed"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="PROBABLE THREONINE SYNTHASE PROTEIN"  
/protein\_id="CAD15030.1"  
/db\_xref="GI:17428343"  
/translation="WQTRSTRGAHPOSFRRILLGLAPDGLYLPQYQVSAEELT  
RMRDLPIADIAFETLRKFATDIPEDDLRALTRTYTEVVCNARNGONTADITPLHL  
GEGAGTSLSLCSNGPTLAKDMAQOLNLPENYLAIDREHAEINLIGATSGDGSAA  
LEYAMRGKRGYRVFMLSPPHRKMSAFQTAQMSLODKNFNIAIEGVPDADIDYAAVN  
DLEAKRQKIGTVASINMARVAVQVYFFGYLLAARKIGDKVFCVPSGNGVACG  
HIAKRMGLPIDKLIVATNENDYIDDEFRTGTGYRYSAAQGYHSSSPMDISKAENFR  
VFVDLGRDPAKLAQLFRVDKRGFSEMTKPEPDRIAEGRFSGRSTHEDRYATID  
VPRYGTGMDTHTADGVKVAARENLYVPPVAVVLETTALPAKFGDTIRALDREPRPAG  
YADIESLPORFEVMPADADRIKGYIAOHTGL"  
4361. 4957  
/gene="RSC1329"  
/note="RS02851"  
4361. 4957  
/gene="RSC1329"  
/function="miscellaneous; hypothetical; global homology"  
/note="Product confidence : hypothetical  
Gene name confidence : hypothetical  
Predicted by Codon\_usage  
Predicted by Homology  
Predicted by Framed"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="CONSERVED HYPOTHETICAL PROTEIN"  
/protein\_id="CAD15031.1"  
/db\_xref="GI:17428344"  
/translation="MAEAFAPSCUHDIRACLCEGALPHEPPVVLADRARVLIQO  
AGRIIVHETGIPWNDSGRDLREMLRIGYDAFYADRIAVPVGFCPPGPGNDLPP  
RPACAPQWHRPRLTAMPDLKLALVGVQADRYADVPGRMLTEVVRARGVPDAEGRV  
LFPDLPSPSPNLMFKNPWFDAQIVRLRAALAAAV"  
5037. 6296  
/gene="moea2"  
/note="RSC1330; RS02852"  
5037. 6296  
/gene="moea2"  
/function="small molecule metabolism; biosynthesis of  
cofactors, carriers; molybdopterin"  
/note="Product confidence : putative  
Gene name confidence : putative  
Predicted by Codon\_usage  
Predicted by Homology  
Predicted by Framed"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="PUTATIVE MOLYBDOPTERIN BIOSYNTHESIS MOEA  
PROTEIN"  
/protein\_id="CAD15032.1"  
/db\_xref="GI:17428345"  
/translation="MSANDAPAYKPSLITLLEALLDOLLAARGLTETETVPTTHANGR  
VLAQAVRSGLVNPPADNTQMDGYAVRADITTAGTRLKVAQRIPACOVGRPLDAGEAA







Score: 265.50 Matches: 195  
 Percent Similarity: 34.28% Conservative: 120  
 Best Local Similarity: 21.22% Mismatches: 348  
 Query Match: 5.62% Indels: 257  
 DB: 6 Gaps: 39

US-09-914-168-2 (1-919) x AX155443 (1-2379)

QY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135  
 DB 46 TCGCTTTGGCATTTGCCGACTTCACCATCCAGATCCGTGTCGAAGGCTTGACGGCT 105  
 QY 136 AsnAsp-----TyrlleProGluTyrGlnGlyGluGlnProAsnSer 149  
 DB 106 ACCGAGCCGAGACCGATTACACTACCTGCCGTCGCAAAAGTCGGGAGACCTACACAC--- 162  
 QY 150 GluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169  
 DB 163 -----GACACACACGCGCACTGCCATCATCAAAAGCCTGTAC 198  
 QY 170 AlaArgLeuPheAsnAspGlyValAlaAsnLysValPro-----ArgLeuLysAlaLys 186  
 DB 199 GCCACCGCTTTCTTGGACGACGTACGAGTCGAAACCTGCCGAGCGGACTTCTGCTGACC 258  
 QY 187 PheTyrGlnSerSerGlnSerGlyLutThrSerAlaIleGlySerSerHisGlnLysThr 206  
 DB 259 GTTATCGAACGCCGCCACCATGGCTCGCTCAACATCACCGGCCCAAAATGCTGCAAAAC 318  
 QY 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluLysP-----LleThrGlnGluSerAla 224  
 DB 319 GAGCGC-----ATCAAGAAAACCTCGAATCGTTCGGGCTGGCCAGTCGCAATAC 369  
 QY 225 MetAsp-----LeuAsnGlySerLleProArgLeuArgGln----- 236  
 DB 370 TTTAATACGCGACACTCAACACGAGTCGCCGCTGAAAGAAAGTAATACCTCGCGGCT 429  
 QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249  
 DB 430 GCGAAACTCAATATCCAAATACGCGCCCAAGTAACCAAAATCGCGCCGACCGCTGCAC 489  
 QY 250 IleAspLeuSerIleIleIleArgAsnSerIleGlyLysValAspValIleIleHisAspLeu 269  
 DB 490 ATGCACTATCAAGATTGACGAGGCAAAATCCGCCAAATACCGCATCGAATTGGA--- 546  
 QY 270 GylGluProValTyrIleAspTyrArgAlaValAlaGluValArgGlyGluGlyAlaAspAsp 289  
 DB 547 GCGAACCAAGTCTATTCGACCGCAAACTGATCGG-----CAG 585  
 QY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309  
 DB 586 ATGCGCTGACCGAAGCGCGCATTTGGACATGGCTGACACGAGCGCGGTTCGACCGC 645  
 QY 310 GylLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329  
 DB 646 CAGAAATTCGCCCAAGACATGAAAGTAACGCACTTACCAACACGCGCTACCTTC 705  
 QY 330 AspGlyLysThrLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349  
 DB 706 GATTTCGCTATCTCGATACCGACATCAACCAACGAAAGCAAAACACGAGCAACCTTC 765  
 QY 350 SerLeuIleTyrAspThrGlyTyrGlnTyrArgPheAspGluValValPhePheThrIle 369  
 DB 766 AAATATCCCTCCACGAAAGCGGAGGATTCGCTGGGCGCAAGTG-----TCGATT 816  
 QY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389  
 DB 817 GAAGGCGACACCAACGAA-----GTCGCCAAGGCGCGAA 849  
 QY 390 LeuGluGlnLeuLeuThrValAsnMetGlyLysLysValTyrAsnLeuGluAlaValAlaArgAla 409  
 DB 850 CTGGAATAACTGCTGACCATTAAGCCCGCAAAATGTTACGAAGCCGACAGATGACCGCC 909  
 QY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429

DB 910 GTTTGGGT----- 918  
 QY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449  
 DB 919 -----GAGATTTCGAACCGCATGGCTGCGAGGCTACCATACAC----- 960  
 QY 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469  
 DB 961 -----GAAATCACGCTACACCCCTGCGCAACCGCGCAACCA 999  
 QY 470 ThrAspGlyIleLeuMetAspLysPro-----IleGlu 481  
 DB 1000 ACCGTCGATTTCGTCTGTCACATCGAACCGGCGGAAATGTACGTCAACGAATCCAC 1059  
 QY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501  
 DB 1060 ATCACCGGCACACAAACCCGCGCAGAA-----GTGCTGGCGCCGCAATTGCGCCCA 1113  
 QY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512  
 DB 1114 ATGGAATCGCGGCTTTACGACACTTCGAAGCTGCAACGCTCCAAAGAGCGCTGACGCTT 1173  
 QY 513 IleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532  
 DB 1174 TTGGGCTACTTCGACACGATCAG-----TTTGAT 1203  
 QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAsnLeuProGlu 552  
 DB 1204 GCCGTCGCCCTTTCGCGGTACGCGCCGCAAA-----GTGATTTTG----- 1242  
 QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572  
 DB 1243 AACATGAGGCTTACCGACACGCTTCAC----- 1269  
 QY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590  
 DB 1270 -----GGCTCGCTGCATTTGAGCGCGGCG 1293  
 QY 591 TTPGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg 610  
 DB 1294 TGGGTTCAGATGATCGGCG-----TTGGTCATGTCCGCGCGGTATCGGACGACACCTG 1347  
 QY 611 AspGlyTyrGlnAlaGlyValGluLeuArgLeuSerGluAspLysLysGlyValLysLeu 630  
 DB 1348 TTCGTTACGGGGAAGTCGCGCCCTGCGCGCTCGCAAGCAAA-----ACCACGCTC 1401  
 QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650  
 DB 1402 AACGCTCGCTGCTGCTTACCGACCGCTACTTCACGCGAGAGCGGGGTCAAGCTGGGCTAC 1461  
 QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670  
 DB 1462 -----GATATTTACGGAAACCTTCGACCGCGCAAGCAAGCATCGACGCTCAAAACA 1515  
 QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTTPAsnArgThrTyrSerGlyLeuArg 690  
 DB 1516 TATTAACCAACCAACCGC-----GGCGGCGGCTAAGATGGG-TATCCCGCT 1562  
 QY 691 TyrArgLeuAspLysLeuLysThrGlnAlaProProGlu-----ThrTyr 705  
 DB 1563 TACCAATATCGACCGCGTCAATTCGGGGCTGGCGGGAACACCTGACCGTCAACACCTA 1622  
 QY 705 PclAspLeuPro-----ValAspPheValAsn-----GlyLys-----ProSerG 719  
 DB 1623 CAACAAAGCACCAACGCTATGCGGACTTTCAGGAATAATGCGGCAAAACCGACGCGGC 1682  
 QY 719 nGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733  
 DB 1683 AGACGGCAGCTTCAAAAGCGCTGTACAAAGCAACCGCTCGGCTGGGCGGCAACAAGAC 1742  
 QY 733 aAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySe 753



Db 910 GTTTGGGT----- 918  
 Qy 430 ProGluArgGluInIleGlnAspGlnValSerPheGluInSerSerSerArg 449  
 Db 919 -----GAGATTGCAACCCGATGGCTGGCAGGCTACGATACAGC----- 960  
 Qy 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469  
 Db 961 -----GAAATCAGGCTACAGCCGCTGGCAGACCCGCGAACCATA 999  
 Qy 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481  
 Db 1000 ACCGTGATATTCCTCTGATGCAACCCGCGGCAAAATCTACGTCAACGAATCCAC 1059  
 Qy 482 PheSerAlaSerAsnLeuIleGlnAspGlyLeuAsnLeuValAlaAlaLysAlaArgHis 501  
 Db 1060 ATCACCAGCAACAAACAAACCCCGCAGAA-----GTCTGCGCCGCGAATTCGCCAA 1113  
 Qy 502 LeuTyraPheMetProAspAsp-----ArgValIleuAla 512  
 Db 1114 ATGGATTCGGCGCTTACGACACCTCCACAGCTGCACAGCGTCCAAAGACGGCTGAGCTT 1173  
 Qy 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532  
 Db 1174 TTGGCTACTTCGACAAACGTACAG-----TTTGAT 1203  
 Qy 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552  
 Db 1204 GCCGTCCCGCTTCCGCTGACGCCGCGACAA-----GTGATTTG----- 1242  
 Qy 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrgInSerLysValPro 572  
 Db 1243 AACATGAGCCTGACCGAACGCTCCAC----- 1269  
 Qy 573 LeuTyraValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590  
 Db 1270 -----GGCTGCTCGATTCGATGAGCGCGCG 1293  
 Qy 591 TrpGlySerAspThrArgLysThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610  
 Db 1294 TGGGTTCAGATACCGGC-----TTGGTATGTCGCCCGCGGATGCGAGGACACCTG 1347  
 Qy 611 AspGlyTyrgInAlaGluAlaGluLeuArgLeuSerGluAspLysGlyValLysLeu 630  
 Db 1348 TTCGGTACGGGCAAGTGGCGCGCTGCGGCTCGCGAAGCAAA-----ACACGCTC 1401  
 Qy 631 TyraIleThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrg 650  
 Db 1402 AACGCTCGCTGTCTTTACCGACCCGCTACTTACGCGACAGCGGGGTCTGGGCTAC 1461  
 Qy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670  
 Db 1462 -----GATATTACGGAAGAACCTTCGACCCGCAACACATCGACAGCGTCAACAA 1515  
 Qy 671 HisGluIleSerArgSerIleIleGlnAsnGlyLysTrpAsnArgThrTyrgSerLeuArg 690  
 Db 1516 TATAAACAACACACCGCC-----GGCGGCGGCTAAGAGATGGG-TATCCCGCT 1562  
 Qy 691 TyraArgLeuAspLysLeuLysThrGlnAlaProProGlu-----ThrTr 705  
 Db 1563 TACCGAATACGACCGCGCTAATTCGGGTGGCGCGCAACACCTGCACCGTCAACACCTTA 1622  
 Qy 705 PglAspLeuPro-----ValAspPheValAsn-----GlyLys--ProSerG 719  
 Db 1623 CAACAAGACCAACCAACGCTATGCCGCTTATCAGAAATACGCGCAAAACCGACGCGC 1662  
 Qy 719 nGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733  
 Db 1683 AGACGCGACGCTTCAAAAGCGCTGTACAAAGGACCGCTGGCGGCGCACACAGAC 1742  
 Qy 733 AspAsnLeuValAsnProMetArgGlyTyraArgGlnArgTyrgSerLeuGluValGlySer 753  
 Db 1743 CGACAGCGCGCTCATGGCCGACGCGCGCTAC----- 1773

Qy 753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrg 773  
 Db 1774 -----CTGACGGCGGTAAATGCGGAATTCGCTCCCGCGACCAATCTCANA 1823  
 Qy 773 rSerPheGlyAspAsnAlaTyrgLysSerAsnArgAlaHisGlnMetThrGly----- 790  
 Db 1824 C-----TACTCCGCCACCCACCAACCAACCTGGTTCTTCC 1859  
 Qy 791 -----GlyIleGlnAlaGlyTyrg 797  
 Db 1860 CTTAAGCAAACTTCACGCTGATGCTCGCGCGCAAGTGGCATTTGGCGGCGCTAC-- 1917  
 Qy 797 eTrpSerAspAsnPheAsnHisValProTyraArgLeuArgPhePheAlaGlyAspAl 817  
 Db 1918 -----GGCAGAACCAAGAAATTCCTTTTGAATACTTCAGCGCGCGCTCGG 1970  
 Qy 817 nSerIleArgGlyTyraIleHisAspSerLeuSerProIleSerAspLysGlyTyrgLeuTh 837  
 Db 1971 TTGGTGGCGGCTACGAAACGGCAGCTCGCGCG-----AAAGTGTATGACGA 2021  
 Qy 837 rGlyGlyGlnValLeuAlaValGly-----ThrAlaGlyTyraAsnTyrgLupheme 854  
 Db 2022 ATACGGCGCAAAATACAGTACGGCGCACAAACCAACGTCCTCGCGCAGCTGCT 2081  
 Qy 854 LysAspLeuArgLeuAlaValPheGlyAspAl 865  
 Db 2082 CTTCGCCGATGCCGCTGGGCAAGCAGCAGCAGCTGCTGAGCTGTGTTGGCGGAC 2141  
 Qy 865 eGly----- 866  
 Db 2142 AGCAGCGCTGGGACGCGCAGACCTATACCGCCCGCAACCGTACCAACATCCGT 2201  
 Qy 867 -----AsnAlaTyraAspLysGlyPheThrAsnAspThrLysIleGlyAlaGly 883  
 Db 2202 TTACTCGGAACCGCGCATTAATCCATTCACACGAAATTCGCTTCGCGCGCGG 2261  
 Qy 883 IGLValArgTrpAlaSerProValIGlyValArgValAspValAlaThrGlyValTyrg 903  
 Db 2262 CGCGGTACCTGGCTGCTGCTTGGTGGTCCGATGAATTCAGCTACGCTACCGCTGAA 2321  
 Qy 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 2322 GAAAAACCGGAAGACGAATTCACACGCTTCCATTCACGCTGCGCAGCAGCTTC 2376  
 RESULT 27  
 AE010493/c 10029 bp DNA linear BCT 01-JUL-2002  
 LOCUS 197 of the complete genome.  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 Fusobacterium nucleatum subsp. nucleatum ATCC 25586.  
 Fusobacterium nucleatum subsp. nucleatum ATCC 25586  
 Bacteria: Fusobacteria: Fusobacterium.  
 1 (bases 1 to 10029)  
 Kapatal, V., Anderson, I., Ivanova, N., Reznik, G., Los, T.,  
 Lykidis, A., Bhattacharya, A., Battman, A., Gardner, W., Grechkin, G.,  
 Zhu, L., Vasileva, O., Chu, L., Kogan, Y., Chaga, O., Goldsman, E.,  
 Bernal, A., Larsen, N., D'Souza, M., Malunas, T., Pusch, G.,  
 Haselkorn, R., Fongstein, M., Kyrpides, N. and Overbeek, R.  
 Genome sequence and analysis of the oral bacterium Fusobacterium  
 nucleatum strain ATCC 25586  
 J. Bacteriol. 184 (7), 2005-2018 (2002)  
 2 (bases 1 to 10029)  
 21886394  
 11889109  
 Kapatal, V., Anderson, I., Ivanova, N., Reznik, G., Los, T.,  
 Lykidis, A., Bhattacharya, A., Battman, A., Gardner, W., Grechkin, G.,  
 Zhu, L., Chu, L., Kogan, Y., Chaga, O., Goldsman, E., Bernal, A.,  
 Larsen, N., D'Souza, M., Malunas, T., Pusch, G., Haselkorn, R.,  
 Fongstein, M., Kyrpides, N. and Overbeek, R.

TITLE Direct Submission  
JOURNAL Submitted (13-FEB-2002) Integrated Genomics, 2201 W. Campbell Park  
Drive, Chicago, IL 60612 USA  
FEATURES  
source  
1. 10029  
/organism="Fusobacterium nucleatum subsp. nucleatum ATCC  
25586"  
/strain="ATCC 25586"  
/sub\_species="nucleatum"  
/db\_xref="ATCC:25586"  
/db\_xref="taxon:190304"  
69. 656  
/gene="FNI907"  
69. 656  
/gene="FNI907"  
/EC\_number="2.1.1.45"  
/codon\_start=1  
/transl\_table=1  
/product="Thymidylate synthase"  
/protein\_id="AAL94006.1"  
/db\_xref="GI:19713161"  
/translation="MKTATYAKKECKIEFEERKSKFTIGYKVPFSKEPEAYIKYN  
LHSDATHNCSAYKINNNGLEFPKVDDEPGSTAGKPMGDIINMEVSNLYATRIE  
GCIKAGAGLVNNAKTAKLAITAEIIDFDIKDILFEIYERLGEYKLAEEYAE  
VIDKSFLEIKYKVKINKNFYDNLENYTFINLDI"  
complement(736. 1809)  
/gene="FNI908"  
complement(736. 1809)  
/EC\_number="3.1.4.46"  
/codon\_start=1  
/transl\_table=1  
/product="Glycerophosphoryl diester phosphodiesterase"  
/protein\_id="AAL94007.1"  
/db\_xref="GI:19713162"  
/translation="MKLSCLVILGLISSALFAANNGITIAHRCASGILPEHTLES  
KALFAQADVLEEDDLMSKGLIYIHDFDGLDLDVAKKFPNRRADGRYVIDFT  
WSELOLEMTENFTTKDKOTAYPMRPLMSKDEFLHFEIEIEFTQGLEKSTGKI  
GIYEIKAPWHEHONGKDIKATLEVKATYGTKSDMYLQTFYNELKRTKTELP  
KMGDKLVLQIAYNDMHETEEKDKGNVNDYDMKEGAMKEIAKYAGVGGVGY  
MLIDDKSKYKGNVYTPMYKDIATTKMELHPYVRKDALPEFTVDVNMAYALLNKA  
ATGVTDFPDGLQFLENQNK"  
complement(1852. 2850)  
/gene="FNI909"  
complement(1852. 2850)  
/EC\_number="2.3.1.-"  
/codon\_start=1  
/transl\_table=1  
/product="UDP-3-O-(3-hydroxymyristoyl) glucosamine  
N-acyltransferase"  
/protein\_id="AAL94008.1"  
/db\_xref="GI:19713163"  
/translation="MEYVRDITILNAEYKGEVEVSKLSFFHSDEKSLTFPADE  
KFLNLSQTKAKVILIPDIDPLEEGGIVAVDSPIVMPKLLHFFSRLLKTEMR  
EDSAKIGENVDIAPVNGVIGNVNNTKIFPNVTIGEGIIIDGIVISVNSIREMR  
ETIKKVRIDNVQIAHNDIIGENCLIIISQVGIAGSTTIGNNVTLAGOVVAGHLEIG  
DNMTIAGOSGPGVNEANKILISCHPLVDHREDMKIVAAKKPELLKRYKALEEK"  
complement(2875. 3348)  
/gene="FNI910"  
complement(2875. 3348)  
/EC\_number="3.1.1.45"  
/codon\_start=1  
/transl\_table=1  
/product="periplasmic protein"  
/protein\_id="AAL94009.1"  
/db\_xref="GI:19713164"  
/translation="MKLLVASVLLATSAFAKGVYVDSKAFPOSETKKAQOSLE  
SOAKVENEAROKREVALOAKDKLTDKAKAFEEKSODFOAFLLNSODKIN  
KEDMAKLEIDVYKAIKKYVADSKYDIFEAELKVGSDITARYLKEEALK"  
complement(1389. 5425)  
/gene="FNI911"

CDS  
complement(3389. 5425)  
/gene="FNI911"  
/codon\_start=1  
/transl\_table=1  
/product="Outer membrane protein"  
/protein\_id="AAL94010.1"  
/db\_xref="GI:19713165"  
/translation="MVNLPKISVEVNNNOVASLTKNTLKIKEGAKFSTALLADFN  
ALKETGFEDVILQPVSDGVRIVDVYEKENVDDLKEGVAINTLRDPDSIVL  
SVKFTGNRKRVTSSELDITQLKAGETISRSRSEDAQRLLATGKFSERDAOVANG  
KMAISFEVNEPPIKSVITITGNNTPTSTINSELTTPKGSQVNNRLEDRDKITLIG  
QAGYTLVNTDMSKIDENGLHISVEGIVRIEYKAVTKOKGRBRRPNDVYKTD  
DRTAELQGGVAVGETGFLGTLSDKSMRNGOQFNGFPKSKNNTTGRALDPEW  
IKOTDRVSWGAVRTSYGDEDSILFEHIDTIGFPTNIGKRLKNTLSQTKVEYIK  
EKHEDEKLRQANNNGKWKYERKKMREITGVDDKWLMSITPYISYDTRNNVLANPSCF  
YKRFQVEAGHAGGYKSGNFGNATLELPTTHGGLKRNNTFAKVVGVAVNNTESQF  
VWGGNSLRGYDGGFFGSKOKLATIENKTLNDIIGLVFADAGRAMKONGRPSYT  
RNSRFGHIGTTAGVIRLNTPIGLPFRFEGWGVGNKMDGKPFYNNQGSF"  
complement(5558. 9085)  
/gene="FNI912"  
complement(5558. 9085)  
/EC\_number="Hypothetical protein"  
/codon\_start=1  
/transl\_table=1  
/product="Hypothetical protein"  
/protein\_id="AAL94011.1"  
/db\_xref="GI:19713166"  
/translation="MDRIDESILNKIAPIREKKLDLKNINIEDIKAVIYSDRRLST  
KTKMPNNSERKGIELDNFNLVYSSKAGKNLSARILTKIPKIPNIALSVNOKNT  
DIIALASPILKNDITPIDINIKAVENOKDIIKANIDSNVDPNVYNNKIKUTLYGN  
KFTINVDYDKRKITNGEGRIPFEIYHTAVYLVAKNNKIQIKELKDKAKNSYFT  
SLEYIGKIERKINGTYNLRKYKKNIEANVYKIDISGYDKRNKNLNTSSQOINDL  
KDSNNMLISLKGQDYVKNOKVANNISVNTKMDVAFKVEVEFENLNAVEGQKIL  
QGVADGSGTITLTPSKDPYIKTGKASIKNSIVNIGDINLNLGKNLADKMDLDR  
VLSKEHLEKYYAKDGLYILGQIDVQVAKIAVAKGATNFERKLPDLADITY  
NAENYADGASITAGDITIDQYIGDLQGVAKLEKTLIDLNKNNKIDLAKIONLIS  
NIDYVKNLVAGNGYIDIKNNYKLDISSSEVSIKTFKINDISNLGDEKRSKLNK  
LJENETKGTINLEINISLESSTAKLILMSNGPINGERIDVGLRSLRNSNPLVNG  
FYDLANISKLEDEDIRSLPYLHKMFHPVPEITIKIGSTELTAEVYGNLI  
IDDAIVYDIPNNYRDFEFLIREOLRKRRYVSTKIDDKOSKTKERKEVEMRLMK  
LMPIDFVYTERKPIILDMDNFIVAPRYGGLYIDLNLNKKGYITGETELKEGYF  
FYGTNFEQVDRALAVENNEPILPEINPIFESEITENDDEYHNTAGVNOLEYS  
SKTAVVGGDLALVFNPNADSHITSYGDNIEITTFKNNLIACQAGOVVGSTRYIK  
RKFDLTKFYIPEVKAITNSDSSVNRLTGTTDNGFPEIYNNIKLEAKDNITKRL  
FKKANVRITIGKDVIKNQMKVDSKYREYVGLKYVDDKQITELGCVTPDKRTD  
PKKDVRFKNYHIGFRFRKRYRDFSEISF"

BASE COUNT 3102 a 1691 c 1027 g 4209 t

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:
311	9.12e-07	10029
329	262.50	162
330	38.08%	111
349	22.59%	265
350	5.35%	179

DB: 1 Gaps: 36

US-09-914-168-2 (1-919) x AE010493 (1-10029)

QY 311 LYSTYRGUThrLysLysAsnLeuIleGlu---AsnAlaSerAlaGluHisGlyTyrPhe 329

Db 5329 AAGTTTTCACACAGAGAGCTTATTAGCTTAAAGCCTTAAAGAAACACAGTTATTTT 5270

QY 330 AspQlyVArGrTirPleuAspArSerValaAspValIleLeuProAspAsnThAlaAspVal 349

Db 5269 GAA-----GATGTAATTTCTTCAA----- 5252

QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArGpHeAsp-----GluValVal 365

Db 5251 CCAAGTTTCTTATGATGCTGAGTGAAGATAGTGTGATGTTGTAAGAAAGAAATGTT 5192

QY 366 PhePheThrIleAspProLys-----ThrsnglnLeuThrAspProAspLys 382  
 Db 5191 GTGATTTATTGAAAGAAAAAGGTCTTGCAATTAACACTTTTACAGAGAAATACAGATATAA 5132  
 QY 383 LeuProVal-----LysArgGluLeuLeuGluGlnLeu 393  
 Db 5131 TCATTTGTCTTCATCAGTAAATTTTACTGGAATTAAGAGATTACTACTACAGAACTT 5072  
 QY 394 LeuThrValAsnMet-----GlyGlnAlaTyrAsnLeuGlnAlaValArgAlaLeu 410  
 Db 5071 TTGAGATTATTCACAAATTAAAGCAGGGAGATTTTTCACAGAGTTCAGAGTTCGAAATCCT 5012  
 QY 411 SerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPhePro 430  
 Db 5011 CAAAGAGATTATTAGGAGCTGGAATTTTCACAGAGTTACACCAGAT----- 4964  
 QY 431 GluArgGluGlnIleGlnAsnAspGln-----ValSerPheGluGlnSerSerSerSer 448  
 Db 4963 -----GCACAGAGTACGAAATGCAAAAATGCTTTATCATTTGAA----- 4925  
 QY 449 ArgThrGluProAlaGlnValAspLysSerThrLeuGluProValIleGlnThrValGlu 468  
 Db 4924 -----GTAGTTGAAAT-----CCAAATGATAAAGTCTAATA 4892  
 QY 469 LeuThrAspGlyIleLeuMetAspLysSerProIle-----GluPheSerAla----- 484  
 Db 4891 ATCAGCAGAAATATATCTATATCCAAAGACTTATATGTCAGAAATGACACTTAACCT 4832  
 QY 485 -----SerAsnLeuIleGlnAspLys-----LeuAsnLeuVal 495  
 Db 4831 GGTTCACCTCCAAATTTATACAAATCTTAAGAGATAGATGATTAATTTTAGGATGTAT 4772  
 QY 496 AlaAlaLysAlaArgHisLeuTyrAspMetProAsp-----AspArgVal 510  
 Db 4771 CAAGCCCAAGATATCTACTTTAGTAATATATACACAGATATGTAAGTGAAGAAATGGAAT 4712  
 QY 511 LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg----- 526  
 Db 4711 TTACATATTTCATATGTAAGGATTTGTAAGAAATGGAAGTGAAGTGAAGTGAAGTGAAG 4652  
 QY 527 -----IleSerAspAlaValSerAlaValAlaArgAlaLeu 538  
 Db 4651 AAACAAAAAGTAAATAGAGAAACACCTAATGATGATGTTTGAAGAACTTAACATATGTT 4592  
 QY 539 LeuProAspLysSerGlu-----AsnGluValIleAspLeuProGluArgThrAla 555  
 Db 4591 ATAGACAGAGAAATAGAAATATACAACTGGAATAATATTATGTTAAAGAAATATGATGCA 4532  
 QY 556 LeuAlaAsnArgLysThrProAlaAspVal-----TyrGlnSerLysLys 570  
 Db 4531 ACAGTTGATTAATGATGATGATGGAATATTTCAAAAATGTTAAATATGAGAACAGATCA 4472  
 QY 571 ValPro-----LeuTyrValPheValAlaSerAspLysProArg 583  
 Db 4471 ATTTCAGAGATCCAGAGAAATGATGATTTGATCTTTAAATGATGAGAGAGAACTCTCT 4412  
 QY 584 AspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPhe 603  
 Db 4411 GAATTACAGGTGGGTTGCTTATGTTCTGAACAGGA-----TTT 4370  
 QY 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlu-----Leu 619  
 Db 4369 TTAGGTACTTATCATTAAGAGATAGTAATCGAGAGAGTAAATAATCAACAATTTGTTTC 4310  
 QY 620 ArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisPro 639  
 Db 4309 ACATTTGAAAAATCAATTAATAATTTACACAGTTTTCATTAAGATTTCTATGATCCTTGG 4250  
 QY 640 LeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThr 659  
 Db 4249 ATTAAGATACAGATAGATGATGATCTTTGGGATGG-----GGAGCTTACAGA 4205

QY 660 AsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSer-----Arg 675  
 Db 4204 ACTAGTTATGGATGATGAGATAGCATACTACTATTCATGAAGAAATAGATCATAGTTCACA 4145  
 QY 676 SerIleIleGlnAsnGly---GlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAsp 694  
 Db 4144 ACTAATATAGGTAAAGAGACTTGTAAGACTTTTACACTAGTTTAGAACAAAGTTGAA 4085  
 QY 695 LysLeuLysThrGln----- 699  
 Db 4084 TACATTAAAGAAACATGACGATGAGAAATTAAGACAAACAATATGTAATGTAATGTAAT 4025  
 QY 700 ---AlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSer 718  
 Db 4024 TTTAAAGAAACAAATTAATAGCAGAGA-----ATGAAAGGTGTAGATACAAATCTCG 3971  
 QY 719 GlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsn 738  
 Db 3970 TTGTGGAGATCTATCTATATCTATATCTATAT-----GACCAAGAAATTAATCTTAAT 3917  
 QY 739 PrometArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSer 758  
 Db 3916 CCTACATCTGATCTGATCTGATGGAATTTCAAGTTGAAGCTGACATGCTGTGGATTAAG 3857  
 QY 759 AspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSer-----PheGly 776  
 Db 3856 TCAGAGAACTTTGGAATGCTACTTGGAAATTTGAGAACATATCATTAAGATTTTAA 3797  
 QY 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyr 796  
 Db 3796 AATATATATATTT-----GCTTACAAAGTTGTGTGAGAGACTACTACT----- 3755  
 QY 797 IleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGlyLysAsp 816  
 Db 3754 -----AACATACAAAGAAAGCAAAATTTCTGGGTGGTGAAGA 3713  
 QY 817 GlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeu 836  
 Db 3712 AATTCATTAAGAGATAT-----GATGAGAGATCTTT 3680  
 QY 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp 856  
 Db 3679 AAAGGAAGCAAAACTT---GTGGCACTATTTGAAGAAATAGAACACACTTAATGATATT 3623  
 QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr-----AspLys 871  
 Db 3622 ATAGAGACTTGTGTATTTCTGATGACGTAGACATGCAACAAATATGAAAGATTCCT 3563  
 QY 872 GlyPheThrAsnAspThr-----LysIleGly-----AlaGlyValGly 884  
 Db 3562 AGTTACACAGAGATATATAGTCTTTTGACATTAACATAGGACACACTGCTGGGTTGGA 3503  
 QY 885 ValArgTTrpAlaSerProValGlyGlnValArgValAspValAla-----ThrGlyVal 902  
 Db 3502 ATCAGACTTAATATCTCAATTCGACCATTAAGATTTGATGCTTTGGTGGCCAGTAGGAAT 3443  
 QY 903 LysGlnGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 3442 AAAATGATGATGATGATGAATGAATTTCTTCAATATATGGACATCAATTC 3392  
 RESULT 28  
 LOCUS NG081959 2379 bp DNA linear BCT 24-JAN-2002  
 DEFINITION Neisseria gonorrhoeae outer membrane protein (omp85) gene, complete cds.  
 ACCESSION U081959  
 VERSION U081959.1 GI:1766041  
 KEYWORDS  
 SOURCE Neisseria gonorrhoeae.  
 ORGANISM Neisseria gonorrhoeae.  
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 REFERENCE 1 (bases 1 to 2379)

AUTHORS Manning, D.S., Reschke, D.K. and Judd, R.C.  
 TITLE Omp85 proteins of *Neisseria gonorrhoeae* and *Neisseria meningitidis* are similar to *Haemophilus influenzae* D-15-49 and *Pasteurella multocida* Oma87  
 JOURNAL Microb. Pathog. 25 (1), 11-21 (1998)  
 MEDLINE 98379445  
 REFERENCE 9705245  
 2 (bases 1 to 2379)  
 AUTHORS Reschke, D.K., Manning, D.S. and Judd, R.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-DEC-1996) Division of Biological Sciences, University of Montana, Health Sciences 104, Missoula, MT 59812-1002, USA  
 FEATURES  
 Location/Qualifiers  
 1..2379  
 /organism="Neisseria gonorrhoeae"  
 /strain="F419"  
 /db\_xref="taxon:485"  
 1..2379  
 /gene="omp85"  
 1..2379  
 /note="omp85"  
 /codon\_start=1  
 /transl\_table=1  
 /product="outer membrane protein"  
 /protein\_id="AACI7600.1"  
 /db\_xref="GI:1766042"  
 /translation="MKIKQIASALMMGLISPLAFADFTIIDIIEVGLQRTPESTVENYLPVAVDVTYNDTHGSAIKISLVATGFEVDVRYETAGDILLTYVCTPGSLITGAKLMONDAIKKMLSEFGLASQYFNQATLNOVAGLKEEYGRGLNTOIPKRYKLARNRVIDITIDEGKSAKITDIEFEGNOYSDPKLMKROSTGSGTGLTMTLSRDRDRKFAADMEKVTDFYONNGYFDFRILIDTIDTNDKTRQTKITVHEGGRFMKVSIEGDTNEVPAKLEKLLTMKPKKTERQMTAVLCEIIONNGSAGYASEISVPLPNAQIKTVPEVLIHERGRKTYNEHITGNNKTRDEVRELRQMESAPYDTSVQPSKERVELCYRDNQFDDPLAGTPPDVLDNMSLTERSTGSLDSAGVQDTGLVMSAGVSDMLFTGKSAALRASRSKRTLPNGSLSTDPDYFTADVSLGDIYGRKAFDRKASTSVKQYKTTAGAGVRMGKIPVTEYDRVNFGLAEHLVNTYKNAKPADEIKQYGTDDAGDSERGLYKGTGVGMGRKTDALMPRTKGYLTGNALETALPGSKLOYSAHNQTPRPPLSKFTLMLGVEGIAGYGRKEIPFEENPFGCGSGVNGYSGTIGRPVYDEYGEKISYSGGNKANVASAILPPMGAKDARTVRISLADAGSVDGRTVTAENGNKSVSEMAHKSTFTEMLKYSAGCAVTMLSPGPMKFTIAYPLKKRPEDEIQRFOPLQCTTP"

BASE COUNT 636 a 752 c 592 g 399 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2, 11e-07 Length: 2379  
 Score: 259.50 Matches: 192  
 Percent Similarity: 33.958 Conservative: 120  
 Best Local Similarity: 20.898 Mismatches: 351  
 Query Match: 5.49% Indels: 257  
 DB: 1 Gaps: 37

US-09-914-168-2 (1-919) x NC081959 (1-2379)

OY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135  
 DB 46 TCCTCCCTTTGGCATTTCGCCGACTTCACCATTCAGACATCCGTGTCAGACGTTTGCAGCT 105  
 OY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149  
 DB 106 ACCGACCCAGACACCGCTATTTCACCTACGTCGCCGTCACAAAGTCGGGACACCTACAC--- 162  
 OY 150 GluAlaValAlaProProThrLeuGluProGluLysProGlyIleuLysArgLeuTyr 169  
 DB 163 -----GACACACACGGGAGTGGCCATCATCAAAAGCCTGTAC 198  
 OY 170 AlaArgLeuPheAsnAspGlyVal-----ArgValLeuAla 177  
 DB 199 GCCACCGCTTTCTTTGACGACGCTACGACGTGAACCTGGGAGCGCTCTTCTGTGACC 258  
 OY 178 -----AsnLysValProArgLeuLysAlaLysPheTyrGlnSer 190  
 DB 259 GTTATCGTATGCCCTACCATCGCTCGCTCAACATCACCGGGCCAAATAATGCTGTGACAGAC 318

OY 191 SerGln----SerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyr 209  
 DB 319 GACGCCATCAAGAAACCCTGCATATCGTTCGGCTGGGCGAGTGGCAATACTTAAATACAG 378  
 OY 210 AlaAsnIleLysAlaIleLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGly 229  
 DB 379 GCGACACTCAACCAAGCAGTGGCGCGCTGAAAGAAATAT-----CTGGGGGCGC 429  
 OY 230 SerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249  
 DB 430 GCGAAACCTCATATCCAAATATCACGCCCAAAAGTAACCAAACTGCCGCAACCGCGTGCAC 489  
 OY 250 IleAspLeuSerIleIleArgAsnSerIleGlyLeuValaValAlaValIleIleHisAspLeu 269  
 DB 490 ATTCACATCAGATTGATGACGAGCGCAATCCGCCAAATATCACCGCAATTCGAATTGAA--- 546  
 OY 270 GlyGluProValTyrIleAspTyrArgAlaValAlaValArgIleGlyGluLysAlaAspAsp 289  
 DB 547 GGCACCAACAGTCTATTCGACCGCAAACTGATGCGG-----CAG 585  
 OY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309  
 DB 586 ATGTCGCTGACCCGAGCGGCGATTTGGACATGCTGACAGAGCGACCGGTTGACCGC 645  
 OY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPhe 329  
 DB 646 CAGAAATTCGCCCAAGACATGAAAGTAACCCACTTACCAAGAAACAGGCTACTTC 705  
 OY 330 AspLysArgTyrLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349  
 DB 706 GATTCCGCTTCCTCGATTCCGATCCGATCCAAACCAACCAAGACAGGACAGCATTG 765  
 OY 350 SerLeuIleTyrAspThrGlyTyrGlnTyrArgPheAspGlyValAlaValPheThrIle 369  
 DB 766 AAATATCCGCTCCAGAAAGCGGACGCTTCCTGCGGCAAGTG-----TTCGATT 816  
 OY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389  
 DB 817 GAAGGGGACACCAACAA-----GTCCCAAGGGCGGAA 849  
 OY 390 LeuGlnGlnLeuLeuThrValaAsnMetGlyAlaTyrAsnLeuGlnAlaValaAlaGala 409  
 DB 850 CTGGAAGAACTGCTGACCATCGAAGCCCGCAATGTGACAGAGCCAGACGATGACCGC 909  
 OY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValaAsnThrGluIleValPhe 429  
 DB 910 GTTTTGGGT----- 918  
 OY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449  
 DB 919 -----GAGATTGAAACCGCATGGGCTCGGACGCTACGATACAGC----- 960  
 OY 450 ThrGluProAlaGlnValaAspGluSerThrLeuGluProValIleGluThrValGluLeu 469  
 DB 961 -----GAAATCAAGCTGACAGCGCTCGACGCGCAAGCGCGGACCAAA 999  
 OY 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481  
 DB 1000 ACCGTGCAATTCGTCTGCTGACATTCGACACCGGCGAGAAAATCTACGTACAGAAATCCAC 1059  
 OY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValaIleAlaLysAlaArgHis 501  
 DB 1060 ATCACCAGGCAACACAAACCGCGACGAA-----GTCTGCGCGCGGGAATTCGGCCAA 1113  
 OY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512  
 DB 1114 ATGGAATCCGCGCTTACGACACCTCCAGGTGCAACCGCTCCAAAGACCGCTGACGCTT 1173  
 OY 513 IleAsnHisAspAspGlyValaAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532  
 DB 1174 TTGGGCTACTTCAACAGCTACAG-----TTTGAT 1203

```

OY 533 AlaValAlaArgAlaIleuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
    ||||| ||| |||||
Db 1204 GCCGTCGCCGCTTCCGGTACGCCGACAA-----GTCGATTTG----- 1242
OY 553 ArgThAlaLeuAlaAsnArgIleProAlaAspValIleGlnSerLysLysValPro 572
    :||| ||| |||
Db 1243 AACATGAGCCTGACCGAAGCTTCCACC----- 1269
OY 573 LeuTyValIleValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
    ||| ||| |||
Db 1270 -----GCTCCGCTCGACTTGAGCGCGG 1293
OY 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610
    ||| ||||| |||||
Db 1294 TGGGTTCAGATACGAGC-----TTGTCATGTCGCCGCCGATGCGAGCAACACTG 1347
OY 611 AspGlyTyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlnAspLysLysGlyValLysLeu 630
    ||| ||| ||||| |||
Db 1348 TTCGGTACGGGCAAGTGGCGCCCTGCGCGCTCGGGAAGCAAA-----ACCACGCTC 1401
OY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
    :||| ||| |||
Db 1402 AACGGTCCTCGCTTCCTTACCGACCCGCTACTTCACGCGACAGCGGGTCACTGCGCTAC 1461
OY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
    :||| ||| |||
Db 1462 -----GATATTACGGAAGACCTTCGACCGCGCAACATCAGTACAGCTCAACAA 1515
OY 671 HisGlnLysSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArg 690
    :||| ||| |||
Db 1516 TATAAACACCAACACCGCC-----GGCGCGCGCTAAGATGG-TATCCCGCT 1562
OY 691 TyrArgLeu-AspLysLeuLysThrGlnAlaProGlu-----ThrTyr 705
    ||||| ||| |||||
Db 1563 TACCGAATACAGACCGGCTCATTCGGCGTGGCGCGGAGAACACTGACCTCAACACTTA 1622
OY 705 PclAspLeuPro-----ValAspPheValAsn-----GlyLys--ProSerG 719
    :||| ||| |||||
Db 1623 CAACAACGACCCAAACGCTATGCGCATTTATCAACAATACGCAAAACCGACGCGC 1682
OY 719 nGluAlaLeuAlaGlyValAlaValHisLysThrVal-----Al 733
    :||| ||| |||||
Db 1683 ACACGGCAGCTTCAAGAGCGCTGCTACAAAGCGACTGCTGGTGGGGCGCAACAGAC 1742
OY 733 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySe 753
    :||| ||| |||||
Db 1743 CGACAGCGCCTTATGCGCGACGCGCGCTAC----- 1773
OY 753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyTyrSerGlyValTyr 773
    :||| ||| |||||
Db 1774 -----CTGACCGCGGTAAATCCCGAATCGCCCTCGCGGACGCAAACTGCAATA 1823
OY 773 rSerPheGlyAspAsnAlaIleTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
    :||| ||| |||
Db 1824 C-----TACTCCGCCACCAACCAACCAACCTGGTCTTCCC 1859
OY 791 -----GlyIleGlnAlaGlyTyrIle 797
    |||||
Db 1860 CTTAAGCAAAACCTTCACGCTGATGCTCGCGCGCGCAAGTGGCATTCGCGCGGCTAC-- 1917
OY 797 eTrpSerAspAsnPheAsnHisValProTyArgLeuArgPhePheAlaGlyLysAspG 817
    :||| ||| |||||
Db 1918 -----GGCAGAACCAAGAAATCCCTTTGAAAACTTACGCGCGCGCTGGG 1970
OY nSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
    ||||| |||||
Db 1971 TTGGGTGGCGGCTACGAAGCGGCGACGCTCGGCGC-----AAATGTGATGACGA 2021
OY 837 rGlyGlnGlnValLeuAlaValGly-----ThrAlaGlnTyrAsnTyrGluPhe 854
    :||| ||| |||||
Db 2022 ATACGGCGCAAAATACGTACGCGCGCAACAAAGCAACAGCATCTCGCGCAAGCTGCT 2081
OY 854 t-----LysAspLeuArgLeuAlaValPheGlyAspIle 865

```

```

Db 2082 CTTCGCCATGCCCGGTGGGAAAGACGACGACCGCTCGGCTGACCTGTTGGCAGCG 2141
OY 865 eGly----- 866
Db 2142 AGGACGCGTGTGGACGCGACAGACTTATCCCGCGCGAAGCGTAAACAAATCGGT 2201
OY 867 -----AsnAlaTyAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVa 883
    ||||| |||||
Db 2202 TTACTCGAAACCGCGCATTAATCACCCTTACACAGAAATTCGCTATTCGCGCGCGG 2261
OY 883 lGlyValArgThrAlaSerProValGlyGlnValArgValAspValAlaThrGlyVally 903
    ||| ||| ||||| |||||
Db 2262 CGCGGTACCTGCGCTCGCTTGGCGCGCATGAATTCATCTACGCTACCCGCTGA 2321
OY 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
    :||| ||| ||||| |||
Db 2322 GAAAAAACCGGAAAGCAATCAACGCTTCCAAATTCAGCTCGGACGAGCTTC 2376

RESULT 29
AX155439
LOCUS AX155439 2394 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 2 from Patent WO0138350.
ACCESSION AX155439
VERSION AX155439.1 GI:14536770
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 2394)
AUTHORS
Giuliani,M.M., Pizza,M., Rappuoli,R. and Holst,J.
TITLE
85kDa neisserial antigen
JOURNAL
Patent: WO 0138350-A 2 31-MAY-2001;
Chiron Spa (IT) ; Statens institutt for Folkehelse (NO)
FEATURES
source
1. 2394
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
BASE COUNT 657 a 743 c 583 g 411 t
ORIGIN

Alignment Scores:
Pred. No.: 2.29e-07 Length: 2394
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
Gaps: 40

US-09-914-168-2 (1-919) x AX155439 (1-2394)
OY 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
    :||| ||||| |||||
Db 46 TCGCCTTTGGCACTTGGCGACTTCACCATCCAAAGACATCCCGGTGGAAGCTTGACAGCT 105
OY 136 AsnAsp-----TyrIleProGluTyrGlnGlnGluGlnProAsnSer 149
    ||||| |||||
Db 106 ACCGAGCGGATACCGTATTCACCTACCTGCGCGTCAAAAGTGGGAGACCTACAC-- 162
OY 150 GluValValAlaProPThrIleGluGluProGluLysProGlyLeuIleLysArgLeuTyr 169
    :||| ||| |||||
Db 163 -----GACACACAGGCAAGTGCATCATTCAAAGCTGTAC 198
OY 170 AlaArgLeuPheAsnAspGlyValAlaLysValPro-----ArgLeuLysAlaLys 186
    ||| ||| |||||
Db 199 GCCACCGGTTTCTTTCAGAGCTACGCGGTGCGAAGCTGGGAGCGGAGCTCTGTCACC 258
OY 187 PheTyGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
    :||| ||| |||||
Db 259 GTTATGCAAGCGCCACCATCGGCTGCTCAACATCAGCGGCGCAAAATGCTGCAAAAC 318
OY 207 GluProTyAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGlnSerAla 224

```

```

Db 319 GACGCC-----ATTAAGAAAAAACCCTCGAATCCTTGGGGCTGGCCGACGTCGCAATAC 369
Oy Metasp-----LeuAnGlySerIleProAlaGluArgGln----- 236
Db 370 TTTAATCAGGGGACACTAATCAGCACTCCGGCCCTGAAAGAAAGAAATACCTCGGGCC 429
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrAsp 249
Db 430 CGCAAACTCAATATCCAAATCAATCAAGTAACTGACCAAACTGCGCCGACCGCTCGAC 489
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
Db 490 ATCGACATCAGCATTTGAGCAGGCAAGTCCGCCAAATACCGACATCGAATTGGAA--- 546
Oy 270 GlyGluProAlaTyrIleAspTyrArgAlaValGluValArgGlyGluValAspAsp 289
Db 547 GGCACCAAGCTATATCCGACCGCAACTGATCCG----- 582
Oy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 583 -----CAATGTCCCTGACCGCAAGCGCGCATTTGGACATGG 618
Oy 307 -----PheHisIleGlyLysTyrGlyThrLysLysAsnLeuIleGlu 320
Db 619 CTGACACGCAACCAATTCACAGCAGAAATTTGCCCAAGATATGAAAAAGTAACC 678
Oy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspVal 340
Db 679 GACTTCTACCAAAATTAACGGGCTACTTCGATTCGATCTCGATACCGACATCCAAACC 738
Oy 341 IleLeuProAlaPheThrAlaAspValSerIleTyrAspThrGlyThrGlnTyrArg 360
Db 739 AACGAAGCAAAACCAACCAAGACCAATCAAAATCACCGTCCAGCAAGCGGAGCTTCCGT 798
Oy 361 PheAspGluValIlePhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 799 TGGGGCAAGTC-----TCCATCGAAGCGCACACCAAGCA----- 834
Oy 381 AspLysLeuProValLysArgGluLeuGluGlnLeuLeuThrValAsnMetGlyGlu 400
Db 835 -----GTCCCAAGACCGCAACTGGAAAAAAGTGTGACCATGAACCGCGCAAA 882
Oy 401 AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 883 TCGTACGAAACCCAGCAGATGACGCGCTTTGGCT----- 918
Oy 421 AsnMetValAsnThrGluIleValPheProGluArgGlnIleGlnAsnAspGlnVal 440
Db 919 -----GACATTCAACACCGCATGGGC 939
Oy 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 940 TCGCGAGGCTACCGCATACAGC-----GAATATCAGCGTA 972
Oy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
Db 973 CAGCGGCTCCGCAACCGTGAACCAAAACCGTCGATTCGTCACATCGAAGCGGGC 1032
Oy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
Db 1033 CGCAAAATCTACGTCACGAATATACATCAACCGGCAACCAAAACCGCGACGAA--- 1089
Oy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
Db 1090 ---CTCGTCCCGCCGTGAATTCAGCAATCCGACCTTACGACACCTCCAAAGCTG 1146
Oy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIle 523
Db 1147 CAACGTTCCAAAGAGCGCGTGAAGCTTTTGGGCTACTTGGCAATGTCTCAG----- 1197
Oy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543

```

```

Db 1198 -----TTTGATGCTGTCCCGCTGGCGGACCGCCGCAAA--- 1233
Oy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 1234 ---GTGATTTG-----AACATGAGTGTGACCGAAGCTTTCACC----- 1269
Oy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 1269 ----- 1269
Oy 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 1270 ---GTTCCCTGGATTGTGACGCGGCTTCAATATCCGGG-----TGGCTATG 1320
Oy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1321 TCCGACAGCGCTTCCCAAGAACACCTGTGCGTACGGGCAAGTGGCGGACATCGCGCC 1380
Oy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1381 TCCAGGAGCAAA-----ACCACGCTTAACGGCTCGCTGTACTGACCCGTAATTG 1434
Oy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlyHisSerThrAsnGly 661
Db 1435 ACGGCACAGCGGCTCAGCTGGCTAC-----GATGTTACGAAAGCCTTCGACCCG 1488
Oy 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 1489 CGCAACAGCATGACAGCATCAACAAATATTAACCAACCAAGCGCA-----GGC 1536
Oy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrAlaGlyLeuAspLysLeu 696
Db 1537 GCAGGATTCGCGATGACGCTGCTTACCGAATACGACGCGGTGAATTCGCTTGGTG 1596
Oy 697 LysThrGlnAlaProProGluThrThrGlnAspLeuPro-----ValAspPheVal 713
Db 1597 GCACAACACCTGACCGGTCACACCTACACAAACCGCCCAACACTATGGCGACTTATC 1656
Oy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
Db 1657 AAGAAATACGCGCAAAACCGACGCGACAGCGGCTTCAAGCTGCTGCTACAAAGT 1716
Oy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 1717 ACCCTGCGCTGGGGCCCAACAAACCGACGCGCTTATGGCCGACGCGGCTAC--- 1773
Oy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 1774 -----CTGACGGGCGTGAACCGCGAATC 1797
Oy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 1798 GCCCTGCTGGCAGCAAACTGCAATAC-----TACTCCGCC 1833
Oy 785 AlaHisGlnMetThrGly----- 790
Db 1834 ACCCAACCAACCAACCTGCTTCCCTGAGCAAAACCTTACGCTGATGCTCGCGGC 1893
Oy 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
Db 1894 GAATCGGCAATTCGGGGCGGCTAC-----GGCAAGAACCAAGAAATCCCTTCTTT 1944
Oy 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
Db 1945 GAAACCTTTCACGCGCGGCTGGTTCGCTGCGCGGATACGAACGCGGACGCTCGGT 2004
Oy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyValLeuAlaValGly----- 845
Db 2005 CCG-----AAAGTCTAGACGAATACGGCAAAAAAATCACTACGCGCGCAACAAA 2055
Oy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 2056 AAAGCAACGCTCCGCGGAGCTCTCTTCCCGATGCCCGCGGCAAAAGACGCGGACCC 2115

```





[illegible]

Accession	Gene	Protein	Length (aa)	Score	Similarity (%)	Best local similarity (%)	Query Match	DB
2056	AAAGCCAAACGCTCTCCGCCGAGCCTCTTCCCATCGCCGCCGAAAGACGGCGCAC		211					
857	LeuArlLeuA1aValAlaPheGlyAspIleGlyValAlaTyrAsp		870					
2116	GTCCGCGTGAACCTGTTTCCGACGACGAGCGGTGGGACGCAAAACCTACGACAC		217					
870	-----		870					
2176	AACAGCAGTTCCGCGACCGCGGAGGCTTCAAAACATTACGGCGCGCAATACCAT		223					
871	--LysGlyPheThrAsnAspThrLysIleGlyValGlyValAlaArgTrpAlaSer		889					
2236	AAATCCACCTTTACCAACGAATGGCTATTCCGCGCGCGCGGCTTACGTGGCTCTGG		229					
890	ProValIleGlyValAlaArgValAspValAlaThrGlyValLysGlu		907					
2296	CCTTGAAGCCGAGAAATTACAGCTACGCTTACCCCTGAGAAAAACCGAAGACGAA		235					
908	ProIleLysLeuHisPhePheIleGlyThrProPhe		919					
2356	ATCCACGCTTCCATTCCATTCACACTCGCAGCAGCTTC		2391					
2394	bp	DNA	Linear	PAT 30-AUG				
AX202489	Sequence 2 from Patent WO0152885.							
AX202489	GI:15392215							
AX202489	Neisseria meningitidis.							
AX202489	Neisseria meningitidis							
AX202489	Bacteria: Proteobacteria; beta subdivision; Neisseriaceae;							
AX202489	Neisseria meningitidis.							
AX202489	1 (bases 1 to 2394)							
AX202489	Pizza, M.C., Rappuoli, R.C. and Giuliani, M.C.							
AX202489	Outer membrane vesicle (Omv) vaccine comprising N. meningitidis							
AX202489	serogroup b outer membrane proteins							
AX202489	Patent: WO 0152885-A 2 26-JUL-2001.							
AX202489	Chiron Spa (IT)							
AX202489	Location/Qualifiers							
AX202489	1. 2394							
AX202489	/organism="Neisseria meningitidis"							
AX202489	/db_xref="taxon:487"							
AX202489	657 a 743 c 583 g 411 t							
AX202489	BASE COUNT							
AX202489	ORIGIN							
AX202489	Alignment Scores:							
AX202489	Pred. No.:							
AX202489	Score:							
AX202489	Percent Similarity:							
AX202489	Best local Similarity:							
AX202489	Query Match:							
AX202489	DB:							
AX202489	US-09-914-168-2 (1-919) x AX202489 (1-2394)							
AX202489	116	ThrpProLeuSerLeuGluGluLeuPheAlaGluGluSerThrGluMetGlyIleAsnPro	135					
AX202489	46	TCGCGCTTTGGGACCTTGGCCGACTTCACCAACAGACATCCGCGCTGAAAGCTTGCAGCGT	105					
AX202489	136	AsnAsp	149					
AX202489	106	ACGAGCGCGAGTACCGATTCACACTACCTCCGCTCAAACTGGCGACACACATCAAC	162					
AX202489	150	GluValValValProPheThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr	169					
AX202489	163	-----GACACACACGCGAGTGCATCATCAAAAGCGCTGAC	198					
AX202489	170	AlaArgLeuPheAsnAspGlyValAlaLysLysValPro	186					
AX202489	199	GCCACGCGGTTCTTTGGTGCAGCATACGCGTCAAACTGGCGACGCGGAGCGTCTGCTGAC	258					

QY 187 PheTyrGlnSerSerGlnSerGlyLutThrSerAlaIleGlySerSerHisGlnLysThr 206  
Db 259 GTTATCGAAGCCGCCACCATTGCTCGCTCAACATCCAGCCGCGCAAAATGCTGCAGAAAC 318  
QY 207 GlnProTyrAlaSerHisIleLysAlaIleLysLysPhe-----IleThrGlnGlnSerAla 224  
Db 319 GAGGCC-----ATTAAAGAAAACCTCGAATCGTTCCGGCGCGCCAGCTCCGCAATAC 369  
QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236  
Db 370 TTTAATATAGCGCGACATCATCATCAGCAGCTCGCCGCTGAAAGAAAGAAATACCTCGGGCC 429  
QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249  
Db 430 GCGAAACTCAATATCCAAATACCGCCCAAGTACCAAACTCGCCCGCAACCGGTCGAC 489  
QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyLysValAspValIleIleHisAspLeu 269  
Db 490 ATCGACATCATCGATTGACGAGGCGCAATCCGCAAAATCACCAGCATCGAATTGGAA--- 546  
QY 270 GlyLysProValTyrIleAspTyrArgAlaValGlyLysLysGlyLysAlaAspAsp 289  
Db 547 GGCACCAACTGTATTCTCCGACCGCAACTGATGGG----- 582  
QY 290 LysAlaPheThrThrValAlaAspGlyValProLeuLeuIleGlyAspVal----- 306  
Db 583 -----CAAAATGCCCTGACCGCAAGCGCGCATTTGGACATCG 618  
QY 307 -----PheHisHisGlyLysTyrGlyLutThrLysAsnLeuIleGlu 320  
Db 619 CTCACACAGCAACCAATTCACGACGACAAATTTGCCCAAGATATGCAAAAGTAAC 678  
QY 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340  
Db 679 GACTGTACCAAAATAGAGGCTACTCTGATTCGATTCGATACCGACATCCAAAC 738  
QY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlyArg 360  
Db 739 AACGAAAGCAAAACCAAGCAGACCATCAAAATCCGTCACGAAGAGCGGCTTCGCT 798  
QY 361 PheAspGlyValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380  
Db 799 TGGGGCAAAATC-----TCCATCGAAGCGCACACCAAGCA----- 834  
QY 381 AspLysLeuProValLysArgGlyLeuLeuGlnGlnLeuLeuThrValAsnMetGlyLys 400  
Db 835 -----GTCCCAAGCGCAGACTGGAATAACTGTCGACCATGAACCGCGGCA 882  
QY 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420  
Db 883 TGGTAGCAAGCCAGCAGATGACCGCGCTTTGGGT----- 918  
QY 421 AsnMetValAsnThrGluIleValPheProGluArgGlnIleGlnAsnAspGlnVal 440  
Db 919 -----GAGATTCAGAACCGCATGGGC 939  
QY 441 SerPheGlnGlnSerSerSerArgTyrGlnProAlaGlnValAspGlySerThrLeu 460  
Db 940 TCGCAGAGCTACGCATACAC-----GAAATGACGGTA 972  
QY 461 GlnProValIleGlyLutThrValGlyLeuThrAspGlyLysLeuMetAspIleSerPro--- 479  
Db 973 CACCGCTGCGCAAGCTGAAACCAAAACCGTCGATTCGTCGACATCGAATCGCGGC 1032  
QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492  
Db 1033 CGGAATATCTACGTCAGCAAAATACATACACCGCGCAACAAACCGCGACGAA--- 1089  
QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508  
Db 1090 ---GTGTCGCGCGGTGAATACGCAAAATGAAATCCGACCTTACGACACCTCCAAAGTGT 1146  
QY 509 -----ArgValLeuAlaIleAsnHisAspArgLysValAsnArgSerIle 523

Db 1147 CAACGTCGAAGACCGCGCTGAGCTTTGGCTACTCTGCAGAAATGCTCAG----- 1197  
QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlnSer 543  
Db 1198 -----TTTGATGCTGTCGCCGCTTCCGCGGACCGCGCGCAAA--- 1233  
QY 544 GluAsnGlyValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563  
Db 1234 -----GTGATTTG-----AACATGAGCTTACCGAAGCTTCCAC 1269  
QY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
Db 1269 ----- 1269  
QY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601  
Db 1270 ---GGTCCCTGGAATTTGAGCGCGGTTGGTTCAGAGATACGGG-----TTGGTCATG 1320  
QY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyValaGlyLeuArgLeu 621  
Db 1321 TCGCAGAGCGTTTCCCAAGACAACTGTTCCGTACGGGCAAGTCCGCGCCACCTCGCGCC 1380  
QY 622 SerGluAspLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
Db 1381 TCCAGAGACCAAA-----ACCAAGCTTAACGCGCTGCTGCTTACTGACCCGCTACTTC 1434  
QY 642 AspGlnLeuArgAlaThrLeuGlnTyrGlnGlnGlnValPheGlyHisSerThrAsnGly 661  
Db 1435 ACGCGACAGCGGGTCAAGCTGGGCTAC-----GATTTAGCAAAACCTTCGACCGC 1488  
QY 662 PheAspLeuSerThrArgThrLeuGlnHisGlyIleSerArgSerIleIleGlnAsnGly 681  
Db 1489 CGCAAGCATGACGACGATCAACATATAAACCACACGGCA-----GCGC 1536  
QY 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696  
Db 1537 GCAGGATCCGCGATGAGCGCTGCTTACCGAATACGACGCGGTGAATTCGGTTGGTG 1596  
QY 697 LysThrGlnAlaProGlnGluThrArgGlnAspLeuPro-----ValAspPheVal 713  
Db 1597 GCAGAACCTGACCGCTGAACACCTTACCAACAAAGCGCCAAACATACTCCGACTTTATC 1656  
QY 714 Asn-----GlyLys---ProSerGlnGlyAlaLeuLeuAlaGlyValAlaValHisLys 730  
Db 1657 AAGAAATACGCGAANAACGAGCGCACAGACGCGACCTTCAAGCGTGGCTGTAACAAGT 1716  
QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744  
Db 1717 ACCGTGCGCTGGGCGGCAACAAACGACGCGGTATGCGCGACGCGCGCTAC--- 1773  
QY 745 GlnArgTyrSerLeuGlnValaGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764  
Db 1774 -----CTGACGCGCGTGAACCGCAAAATC 1797  
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
Db 1798 GCCGTGCTGGCAGCAAACTGCAATAC-----TACTCCGCGC 1833  
QY 785 AlaHisGlnMetThrGly----- 790  
Db 1834 ACCGCAACAACCACTGCTTCCCTGACCAAAACCTTCACGCTGCTCGCGCGC 1893  
QY 791 -----GlyIleGlnAlaGlyTyrIleThrPheAspAsnPheAsnHisValProArg 808  
Db 1894 GAAGTCGCGCATTCGGCGCGGCTAC-----GCCAGAACCAAAATCCCTTCTTT 1944  
QY 809 LeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrGlnAlaAspSerLeuSer 828  
Db 1945 GAAACATCTTACGCGCGCGCTGGTTCGCGTGGCGATACGAAAGCGGACGCTCGGT 2004  
QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGly----- 845



Db 1090 ---GTCTGCGCCCGCAATTCGCCAATGCCGCTTACGACACCTCCAAAGCTG 1146  
 Qy 509 -----ArgValIleuAlaIleAsnHisAspGlyValAsnArgSerIle 523  
 Db 1147 CAAGCTCCAAAGACGGCTGAGCTTTGGCTACTTCGACACGCTACG----- 1197  
 Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspIleuSer 543  
 Db 1198 -----TTTGATGCCGCTCCGCTTGGCGGACACACCGCAACAA--- 1233  
 Qy 544 GluAsnGluValIleAspLeuProGlyArgThrAlaLeuAlaAsnArgLysThrProAla 563  
 Db 1234 -----GTGCAATTTG-----AACATGACCTGACCGAAGCTTCCACGC----- 1269  
 Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
 Db 1269 ----- 1269  
 Qy 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601  
 Db 1270 ---GGCTCGCTCGACTGAGCGGGGCTGGGTACAGATACCGGC-----CTGGTCAATG 1320  
 Qy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlyLeuArgLeu 621  
 Db 1321 TCCGCGACGGCTTTCGCAAGACAACTGTCGTACGCGCAAGTCGCGCCCTCGCGCGCC 1380  
 Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
 Db 1381 TCAGAGAACAA-----ACGACGCTCAACGGCTCGCTGCTTACGACCCGCTACTTC 1434  
 Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661  
 Db 1435 ACGCGACAGCGGGCTACGCTGGGCTAC-----GATGTTACGGGAAAGCCTTCACACCG 1488  
 Qy 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681  
 Db 1489 CGCAAGACATGACCATCAGCATCAACATATTAACCCACGCGCA-----GGC 1536  
 Qy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696  
 Db 1537 GCAGGATCCGATGAGCGCTGCTGTACCGAATGACGACCGGTGAATTCGTTGGTG 1596  
 Qy 697 LysThrGlnAlaProProGlyThrTyrGlnAspLeuPro-----ValAspPheVal 713  
 Db 1597 GCAGAACACTGACCTGACACCTACAAACAAAGCGCCAAACACTATGCCACTTTATC 1656  
 Qy 714 Asn-----GlyLys-----ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730  
 Db 1657 AAGAAATACGGCAAAACCGACGCGACACGCGCACTTCAAGGCTGCTGACAAAGGT 1716  
 Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744  
 Db 1717 ACCGTGGGCTGGGGCGGCAACAAACGACAGCGGTTATGCGGACCGCGGCTAC--- 1773  
 Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetalIle 764  
 Db 1774 -----CTGACGGGCGGTGAACCGCAAAATC 1797  
 Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
 Db 1798 GCGCTGCGCGGACGAAACCTGCATAC-----TACTCGCGC 1833  
 Qy 785 AlanIsglnMetThrGly----- 790  
 Db 1834 ACCGACACCAAACTGCTTCTCCCTTAAGCAAACTTCACGCTGATCGCGCGC 1893  
 Qy 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808  
 Db 1894 GAAGTCGCAATTCGCGCGGCTAC-----GGCAGAACCAAGAAATCCCTTCCTT 1944  
 Qy 809 LeuArgPhePheAlaGlyGlyAspIleSerIleArgGlyTyrAlaHisAspSerLeuSer 828  
 Db 1945 GAAAACTTCTACGGCGGCGCTGGTTCGTCGCGGATACGAAGCGCGCGCTCGGT 2004

Qy 829 ProIleSerAspLysGlyTyrIleThrGlnGlyGlnValIleuAlaValGly----- 845  
 Db 2005 CCG-----AAAGTATACCAATACGCGCAAAATACGCTACGCGCGCAACAA 2055  
 Qy 846 ThrAlaGlyTyrAsnTyrGlnPheMet-----LysAsp 856  
 Db 2056 AAAGCCACAGCTCCGCGGAGCTGCTTCCGATGCCCGGCGGAAAGACGCGCGCAC 2115  
 Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870  
 Db 2116 GTCCGCTGACGCTGTTGCGGACGACGACGCGGTGGAGCGCAAAACCTACGACGAC 2175  
 Qy 870 ----- 870  
 Db 2176 AACAGCACTCCGCGACCGCGCGCAAGGTTCAAAACATTCAGCGCGCGCAATCCAT 2235  
 Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrPheAsp 889  
 Db 2236 AAATCCACCTTATACCAAGCAATTCGCTATTCGCGCGGCGCGGCTACTGCGCTCG 2295  
 Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907  
 Db 2296 CCTTAAAGCCGCGATGAATTCAGCTACGCGCTACCGCGTGAAGAAACCGGAAGACGAA 2355  
 Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 2356 ATCCAACGCTTCATTCACATCTCCACCTCGCACGACGCTTC 2391  
 RESULT 33  
 AE002375/c 13538 bp DNA linear BCT 25-MAY-2000  
 LOCUS Neisseria meningitidis serogroup B strain MC58 section 17 of 206 of  
 DEFINITION the complete genome.  
 ACCESSION AE002375 AE002098  
 VERSION AE002375.1 GI:7225394  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Neisseria meningitidis MC58.  
 Neisseria meningitidis MC58  
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
 Neisseria.  
 REFERENCE 1 (bases 1 to 13538)  
 AUTHORS Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,  
 Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,  
 Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,  
 Hickey, E.K., Nelson, M.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,  
 Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,  
 Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,  
 Qiu, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V.,  
 Pizzza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,  
 Rappuoli, R. and Venter, J.C.  
 Complete genome sequence of *Neisseria meningitidis* serogroup B  
 strain MC58  
 JOURNAL Science 287 (5459), 1809-1815 (2000)  
 MEDLINE 20175755  
 PUBMED 10710307  
 REFERENCE 2 (bases 1 to 13538)  
 AUTHORS Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,  
 Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,  
 Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,  
 Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,  
 Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,  
 Qiu, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V.,  
 Pizzza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,  
 Rappuoli, R. and Venter, J.C.  
 Direct Submission  
 Submitted (17-MAR-2000) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 TITLE JOURNAL  
 FEATURES  
 .source  
 1. 13538  
 /organism="Neisseria meningitidis MC58"  
 /strain="MC58"

```
gene
CDS
/db_xref="taxon:122586"
/notes="serogroup: B"
/complement(121. .1377)
/gene="NMB0176"
/complement(121. .1377)
/notes="NMB0176"
/notes="similar to GB:U02948 SP:P29011 PID:145703 GB:U00096
PID:1651591 percent identity: 76.98; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="D-amino acid dehydrogenase, small subunit"
/protein_id="AAF40633.1"
/db_xref="GI:7225395"
/transl_table="MKVLVAGVAGVSSAMYLAEAGEVTVIDRAGVAMETSPANA
GOLSYGTTTMAAPGIPITKALKMLFSSHPPLPSPDGLQIEMQMIONCTAAHYO
TNKRWVRISEYSEMRERREAOFGNMEGRRKGTLOIFROTREVEAKDIALEHT
GVYRRLKPECAEFEPALARVAKIAGGHLADATGDCRLTENLYKLCQKGVAF
HFNQNTSRIDNCRIRKTVETETGRFADAVCALGCFSTVLAQDLNDLPYVKG
SLPLVYNSDGAIPSTVLDESIVALTFRPNRIRVGMELSGEIKLPEKRETTAL
VVDLPFEGGDLQASAFMSGRLRPTPDSPPLIGTRPRDNLFLNTGHTGLWTGLSA
KLADIVSGKDTERRSDLSLRYOA"
/complement(1456. .2847)
/gene="NMB0177"
/complement(1456. .2847)
/notes="similar to GB:AE000511 PID:2314080 percent
identity: 69.11; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="sodium/alanine symporter, putative"
/protein_id="AAF40634.1"
/db_xref="GI:7225396"
/transl_table="MQVFLDNPKAFETVSGWVWGPIMLMLVGTILTVLLKGF
TMVLVLAQAFMPKPKSGEGHEGDISHFALMTLSATIGTGNAGVATVVTGCP
GAVFMWMTAIFGMATKRGEGVLAAYRVNNSGEMSGPMYIEKGLGKNNKMA
FALFETFAFGSGSVOSNSVAOVOTSGFSGPAGYITLTVLAVYVIGKIKTA
ASPIVPAWVYVLGSLTIAINSNDMPALVLTIPSDASAAVAGATGIVTIGVA
RGVSENEGMSAPITAAAKDHPRLQALVYMTGTFIDTIVVCSITGILVWGLGA
GCFPVREYSGALTTVTQKMLPGIGNIVTIGLIFVYSPILGWCYGECAVYVF
GEKFACTLRVGVSSVMLGTVLSLDLWLASDTFNGLMALPNLIALMLAKYIVETR
DRQKITTNGELPH"
/complement(3204. .3980)
/gene="NMB0178"
/complement(3204. .3980)
/notes="similar to PID:1718489 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="acyl - (acyl-carrier-protein) --UDP-N-
acetylglucosamine O-acyltransferase"
/protein_id="AAF40635.1"
/db_xref="GI:7225397"
/transl_table="MTLIHPTAVIDPKAELDSGVKGVATVIGPNVQIGANTEIGP
VINGTISGENNRIFOFASLGEIPQDKRYDEPTKLIGNGTIREPTFNIGTGTGI
GDTRIQDMNIMAYCHLAHDCVIGNHTIFANNASLAGVTTIGDYYVGLGTYLVOFGR
IGDYATWFAAGVHKVDVDPYFMAISGVRAPAGLSEGRNRGFTAEQISAVKDYKTL
YHRGIPFEAKDILIRAEQALAVDFDFAQSAKGLR"
/complement(4067. .4516)
/gene="NMB0179"
/complement(4067. .4516)
/notes="similar to PID:1718488 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product=" (3R)-hydroxymyristoyl - (acyl carrier protein)
dehydratase"
/protein_id="AAF40636.1"
/db_xref="GI:7225398"
/transl_table="MDVQLPIEAKDIOKLIPHRPFIQLDRITAFEPKKTITAIKNVT
INEPQEGHPDLPVMPGVLIIEPMAQCTLLASEGGRKEFFFAIDEAREFR
QVLPQDQLEVEVELTISRIGIKGFENNAKVDGQVAVAEALIMAKRRV"
/complement(4450. .5596)
/gene="NMB0180"
/complement(4450. .5596)
/notes="NMB0180"
/notes="similar to PID:1718487 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="UDP-3-O-(3-hydroxymyristoyl)-glucosamine
N-acyltransferase"
/protein_id="AAF40637.1"
/db_xref="GI:7225399"
/transl_table="MIPATYTLQITARLGMGRGDISVTAVRPLADAQAEHIFLA
NPKYKAEVHDSAGAVYSAKADGEGENLIVADPPYLYFPAKVRILSPYKARGGI
HPYAVPEGATVPTSCIEIGANYIGANTVIGSGCRILAAVVOHCKIGDEVILPNA
VYVGCTLGRVREIHSAGAVIGADGELAFADDSKRIPTQAGVTTIGDDVEIGSNTNID
RGMSPSTTGNGTKIDNOYQIGHCKTGSHTYIAAKTGISGVTTIGSTCIGGVGTV
GHEIADKTTIGGCTSVTHSITESGHLAGIFPMSTHEWARNNAVYIHRLEPMNRK
TIEQQLSDAGQDSK"
/complement(5629. .6129)
/gene="NMB0181"
/complement(5629. .6129)
/notes="similar to GB:M34854 SP:P31519 PID:155479 percent
identity: 46.88; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein OmpB, putative"
/protein_id="AAF40638.1"
/db_xref="GI:7225400"
/transl_table="WTRLRAPFAALLDCTAGAHADTFQKIGINTERIYLESQA
KRIQTLTSESRARDEQLQRELDLEROLAEKLNARKAAEEMKRLVAFRA
KOAEDEBYNLRNREFASLQONANRVIKAKOEGYVILQNNIYVNTQYDVDSVI
KEMNR"
/complement(6195. .8588)
/gene="NMB0182"
/complement(6195. .8588)
/notes="NMB0182"
/notes="similar to GP:2460281 percent identity: 99.87;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein Omp85"
/protein_id="AAF40639.1"
/db_xref="GI:7225401"
/transl_table="MKLQIASALMMLGISFLADFTQDIRVEGLQRTPESTVENY
LPAVGDYNDTHSALIKSLYATGFDDVAVETADQQLLTVIERPTIGSLNTGAK
MLQNDATKRNLESGLAOSQYFNQATLNOAVAGLEKLEYLGKLNIOITPKVRLAN
RVDIDITIDEKSAKINTDIEEGNOQVSDRKLQMSLITGEGITMTLIRSNQFEOKF
AODMEKVTDFQNNNGYEDFERLIDTIDTQNEKTKQCTKITVHEGGRMKVSEBGT
NEVPAEELKELTKMKPKKWEERQOMTAVLGEIQRMKSAGATASEISQVLPAAET
VDVLIHIEPKKITVNEIHTGKNKTIDEVVRRLRQMESAPDTSKLSQSKREVELL
CYFENVPDAVPLGTPDKVDLNNKTIESTGSLDSAGWODTGLWMSAGVQDNELF
TGGSAALRNSRSYTTLNGSLSPFDPTFTADGVSLGADVGAFFDPKASTSJKOYT
TTPAAGIRMSVPVTEYRVNVEGLVAEHLITVYAKARHADFLKTKGKTDGNDGSK
GWLKGTGVGWRNKTDLSALMPTRGYLGNVAELTAPSKLOXYSAHNNQWPPPLST
FTLMLGGEVGIAGYGTGRTKEIPREPNYGGCLSVKRGESGTIGPKYIDYIGKISIG
GNKKNVSAEILPMPGAKDARTVRLSLFADAGSVWDGKTYDNNSSATGGRVQNTYG
AGNTHKSTFTNELRISAGAVTWLSPGLPKFSAVPLPKKRPDEIORFOGLGTFF"
/complement(8645. .9985)
/gene="NMB0183"
/complement(8645. .9985)
/notes="conserved hypothetical protein; identified by
Climmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40640.1"
/db_xref="GI:7225402"
```

/translation="MHTLLAFAPALLIVSLHSEEDHIVPALDGVKVVRFSEVGGKGRF  
 FRRKSGDIEWCLAPRIPGCVKMDRTREGSEVDELPAPRKHQPKARRIAVAAGPLTL  
 NIALVALILGTSFSFGVTELRPIYVSTVEPTITAAABGQSDGRQIOSVNGVFAVMGNSA  
 QVEIYLVNACKVAAGVQVDTASGADVFRITDAAGPEAKLAKNGOYIGLMEKFTITVAAG  
 GVEEYGVSPAEKAGKLPRGRLTAAODKPRIASQEWANILTRQSPGKRLTLNVERAQOTHTPL  
 ADIRDTYEOSDHTLIGRVGRLPPQDPRAMDQOIRRSYSPYVFAFGMGMEKTVESHSTPT  
 TLKFGKGLISGNASVSHISGPTLADIADISQWAEIGLOSYLEFLIALVSIQCLVNLPLP  
 PVLDCGHLVYFYAEWIRGKPLGERVONIGLRFGLALMLMLMAVAVFNDVTRLLG"  
 complement(10020. .11204)  
 /gene="NM00184"  
 /gene=complement(10020. .11204)  
 /gene="NM00184"  
 /gene=complement(10020. .11204)  
 /gene="NM00184"

Alignment Scores:	
Pred. No.:	2.25e-06
Score:	259.00
Percent Similarity:	33.48%
Best Local Similarity:	20.69%
Query Match:	5.48%
DB:	1
Gaps:	40
length:	13538
Matches:	191
Conservative:	121
Mismatches:	342
Indels:	278
Gaps:	40

US-09-914-168-2 (1-919) x AE002375 (1-13538)

Oy	116	ThpIleuSerIleuGIuIleuPheAlaInGIuInSerThrGIuMeGIyIleAsnPro	133
Db	8543	TCGGCTTTGGCACTTGCCGCATTCACCATTCAAGACATCCGGCTGGAAGCTTGACGCT	8484
Oy	136	AsnAsp-----TyrIleProGIuTyrGIuGIuInProAsnSer	149
Db	8483	ACCGAGCCGAGTACCTGATTCACTACGTCCCGTCAAAAGTCGGGAGACCTACAAAC---	8427
Oy	150	GIuValValValProProThrIleuGIuInProGIuInLysProGIuIleuIleLysArgIleuTyr	169
Db	8426	-----GACACACAGCGAGTGCATCATCAAAAGCCCTGTAC	8391
Oy	170	AlaArgLeuPheAsnAspGIuValAlaInLysValPro-----ArgIleuLysAlaLys	186
Db	8390	GCCACCGGTTCTTGTAGACAGCATCGCGTCGAAGCTCGGAGCGGACGCTCTCTGTAC	8331
Oy	187	PheTyrGIuInSerSerGIuInSerGIyGIuThrSerAlaIleGIySerSerIleuLysThr	206
Db	8330	GTTATTCGACACGCCACCATCGCGCTGCATCATCACCGCGCAAAATGCTGCACAAAC	8271
Oy	207	GIuProTyrAlaAsnIleLysAlaIleuGIuAsp-----IleThrGIuInSerAla	224
Db	8270	GAGCCC-----ATTAAAGAAAACCTCGATCGTTGGGCTGGCGCATGCGCAATAC	8220
Oy	225	MetAsp-----LeuAsnGIySerIleProArgLeuArgIn-----	236
Db	8219	TTTATTCAGCGGACATCATTCAGCAGCATGCGCGGCTTAAGAAAGATATCTCGGGCGC	8160
Oy	237	-----ThrAlaLeuValAlaIleArgAlaValGIyTyrTyrAsp	249
Db	8159	GGCAAACTCATTCACAAATCACGCGCCCAAGTACCAAACTGCGCGCAACCGCGTCGAC	8100
Oy	250	IleAspLeuSerIleIleArgAsnSerIleGIyGIuValAspValIleIleHisAspLeu	269
Db	8099	ATCGACATCATCGATTGACGAGGGCAAAATCCGCCAAATACCGACATCAATTTAA---	8043
Oy	270	GIyGIuProValTyrIleAspTyrArgAlaValGIuValArgGIyGIuGIuLysAspAsp	289
Db	8042	GGCAACCAAGCTATTTCGCACGCCCAAACTGATCGG-----	8007
Oy	290	LysAlaIlePheThrThrValAlaAspGIuValProLeuLeuIleGIyAspVal-----	306
Db	8006	-----CAATGTCCCTACCGAAGGCGGCGCTTTGGACATCG	7971
Oy	307	-----PheHisIleGIyLysTyrGIuThrLysLysAsnLeuIleGIu	320
Db	7970	CTGACACGACCAACCAATTCACAGCAGAAATTTGGCCCAAGATGTGGAATAACTATACC	7911
Oy	321	AsnAlaSerAlaGIuHisGIyTyrPheAspGIyArgTyrLeuAspArgSerValAspVal	340

D	7910	GACTTCTACCAAAATTAACGGCTACTTCGATTCCTCCGATTCCTCGATACCGACATCCAAAC	7851
Q	341	ILEuPProAspAsnThrAlaAspValSerLeuILEtYrAspThrGLYrGlnTYrArg	360
D	7850	AACGAGCAAAACCAACGAGACACCATCAAAATCACCGCTCCAGAAAGCGGAGCTTCCGT	7791
Q	361	PheAspGluValValPhePheThrILEAspProLYrSthAsnGlnLeuThrThrAspPro	380
D	7790	TGGGGCAAAAGTC-----TCCATTCGAAGCGACACCAACGAA-----	7755
Q	381	AspLYsLeuProVALYsArgGluLeuLeuGlnGlnLeuThrValAsnMetGLYlu	400
D	7754	-----GTCGCCAAAGCCGAACCTGGAATAAACTCTACCATCAAGCCGGCAAA	7707
Q	401	AlaTYrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuILEalaThrArgTYrPhe	420
D	7706	TGTATCGAAGCCGACGAGATGACCCGCTTTGGT-----	7671
Q	421	AsnMetValAsnThrGluILEValPheProGluGlnGlnILEGlnAsnAspGlnVal	440
D	7670	-----GAAATTTCAGAACCGCATGGGC	7650
Q	441	SerPheGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu	460
D	7649	TCCGCACGGCTACGCATACAC-----GAAATCACGCTA	7617
Q	461	GluProVALILEglnThrValGluLeuThrAspGLYILEuMetAspILESerPro---	479
D	7616	CAAGCGGCGCGAAGCGCTGAACCAAAACCGCTGATTTGCTCTCACATCGAACCGGCG	7557
Q	480	-----ILEglnPheSerAlaSerAsnLeuILEglnAspLYsLeu	492
D	7556	CGGAAATCTACGTCACAGAAATATCACATCACCGGCACAAACAAACCCGCGACGAA---	7500
Q	493	AsnLeuValAlaAlaLYsAlaArgHisILETYrAspMetProAspP---	508
D	7499	-----GTCCGTCGCCGCTGAATTACGCCCAATGGAAATCCGCACTTACGACACCTCCAAAGTG	7443
Q	509	-----ArgValILEuAlaILEasnHisAspAspGLYValAsnArgSerILE	523
D	7442	CAACGTTCCAAAGAGCGCGTCGAGCTTTGGCTACTTCGACAAATGTCCAG-----	7392
Q	524	LeuGLYrArgILESerAspAlaValSerAlaValAlaArgAlaILEuProAspGluSer	543
D	7391	-----TTTGATTCGCTGCCCGCTTGGCGCGACCGCCGACAA-----	7356
Q	544	GluAsnGluValILEAspLeuProGluArgThrAlaLeuAlaAsnArgLYrSthProAla	563
D	7355	-----GTCCGATTTGG-----AACATGAGCTTACCGGAAGCTTCCACC-----	7320
Q	564	AspVALYrGlnSerLYsValProLeuTYrValPheValAlaSerAspLYrProArg	583
D	7320	-----	7320
Q	584	AspGLYrGlnILEgLYLeu-----GLYrTPGLYSerAspThrGLYrThrArgLeuValThr	601
D	7319	---GGTTCCTCCGTGATTGTAGACGGCGGGTGTGGGTTCAAGATACCGGG-----TTGGTCTATG	7269
Q	602	LYsPheGlnHisAsnLeuILEasnArgAspGLYTYrGlnAlaGLYValAGluLeuArgLeu	621
D	7268	TCCCGCAGCGCTTTCCCAAGACACCTGTTGGGTAGCGGCAAGCTCGCCCATCGCGCC	7209
Q	622	SerGluAspLYsLYsGLYValLYsLeuTYrAlaThrLYsProLeuSerHisProLeuAsn	641
D	7208	TCCAGAGACCAA-----ACCACGCTTAAACGGCTCCCTGCTGTTATCTACACCCGCTACTTC	7155
Q	642	AspGlnLeuArgAlaThrLeuGLYrTYrGlnGlnGlnValPheGLYHisSerThrAsnGLY	661
D	7154	ACGCGACAGCGGGGTACACCTGGGCTAC-----GATGTTTACGGAAGAAAGCTTCGACCGG	7101
Q	662	PheAspLeuSerThrArgThrLeuGlnHisGluILESerArgSerILEglnAsnGLY	681
D	7100	CGCAAAAGCATCGACCATCAAAACATATAAAAAACCAACCCACGAGCA-----GGC	7053

[illegible]

RESULT 34  
AC020884\_1/c  
WPCOMMENT  
Sequence split into 4 fragments LOCUS AC020884  
Fragment Name Position End Accession AC020884

Fragment Name	Begin	End
AC020884_0	1	110000
AC020884_1	100001	210000
AC020884_2	200001	310000

```

AC020884_3      300001      373038
Continuation (2 of 4) of AC020884 from base 100001 (AC020884 Mus musculus clone RP23-
Alignment Scores:
Pred. No.:      3.57e-05      Length:      110000
Score:          259.00      Matches:      120
Percent Similarity: 35.09%      Conservative:  86
Best Local Similarity: 20.44%      Mismatch:     192
Query Match:    5.48%      Indels:       190
DB:             2      Gaps:         17
US-09-914-168-2 (1-919) x AC020884_1 (1-110000)

```

QY	272	ProValTYrTLeAsPTyRArgAlaValAlaGluValAArgGluGluGluAlaAspLysAla	291
Db	42289	CCGGTGAATTAATGGCGGACCGAGTGTGGTATTGGCGGCGGCGCGGACCGATTAAGC	42230
QY	292	PheThrThrValAlaAspGluValProLeuLeuLeuIleGlyAspValPheHisGlyLys	311
Db	42229	TATTTGAATATGCTCGAATACGCGCCG---GCTATTGGACAGTACTGAACAGCGGAT	42173
QY	312	TyrGluThrLysAsnLeuIleGluAsnLaserAlaGluHisGlyTyrPheAspGly	331
Db	42172	TATGGAATTTTCAAAAGTCCCTTAACGACGATTCGCGTTGGCTTAAGGTTATTGGATGAC	42113
QY	332	ArgThrPleuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLys	351
Db	42112	GAATTTACCAAGCGCAGCATTCGGCATTCGCTCGGCTTCGATTAAGCCTTCGGATATT	42053
QY	352	IleTyrrAspThrGlyThrGluTyrArgPheAspGluValAlaPhePheThrIleAspPro	371
Db	42052	GATTATTAACAGTGGGAGCAAGCTTACCGCTTTGGCATGTGACCTTT-----	42008
QY	372	LysThrAsnGluLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu	391
Db	42007	-----GAAGATCACAAATCCGCGCATGAATACCTTCGA	41975
QY	392	GluLeuLeuThrValAsnMetGlyGluAlaTyrrAsnLeuGlnAlaValAlaArgAlaLeuSer	411
Db	41974	AATCGGAGCCGTTTAAAGCGGAGTAGTACGAATCGAAGATCGCGCAACTGAC	41915
QY	412	AsnAspLeuIleAlaThrArgTyrPheAsnMetAlaSnThrGluIleValAlaPheProGlu	431
Db	41914	CGCGACATCTTGCTACCGCGGTGGTTAATCCGGTG-----GTGGTGGCTCCACAA	41864
QY	432	ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerSerArgTyrGlu	451
Db	41863	-----TTTATTAAGCGCGCAAGCAAGGATTTA	41834
QY	452	ProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp	471
Db	41833	CCA-----TTGACG---	41825
QY	472	GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys	491
Db	41824	GGCGTG-----GTTTCGCGG---	41810
QY	492	LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrTraspMetProAspAspArgValLeu	511
Db	41810	-----	41810
QY	512	AlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGluArgIleSerAspAlaVal	531
Db	41810	-----	41810
QY	532	SerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValAlaAspLeuPro	551
Db	41809	-----CGAACGAGAAACACCATC-----	41792
QY	552	GluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysVal	571
Db	41792	-----	41792



QY 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyr 591  
 Db 41791 -----GAAACCGGGCGGTAC 41774  
 QY 592 GlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAspLeuIleAsnArgAsp 611  
 Db 41773 TCTACGGACGGTGGACCGCGGTGAACCGAGCTGGMAAACCCGTGGATGACCTTAT 41714  
 QY 612 GlyTyrGlnAlaGlyAlaGlyLeuArgLeuSerGluAspLysGlyValLysLeu 630  
 Db 41713 GGTCAACATCTGCACACACAGCTAGTATTTCCGGCCGGACACACACCTCCACTTCAC 41654  
 QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650  
 Db 41653 TATAAATGCGCTGCTGCAAAATCCTGCAACATATTATTGTCGACGGCGGTTT 41594  
 QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670  
 Db 41593 AAGCGCACT-----GACCTGAACGATACCGAATCTGACTGCACCTACGCTG 41549  
 QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArg 690  
 Db 41548 ---GTGGCTTCTCGTACTGGGATCTCTCCACGGCTGGACACGTCCTTAACCTCGCC 41492  
 QY 691 TyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProVal 710  
 Db 41491 TGGAGTCTGCAC----- 41480  
 QY 711 AspheValAsnGlyLysProSerGlnGluAlaLeu-----AlaGlyValAlaVal 728  
 Db 41479 CACTTTACTAGAGGTAATTAACAATACACAGATGCTGTTTATCTGGGTGATGATT 41420  
 QY 729 HisLysThrValAlaAspAsnLeuValAsnProMetArgLysTyrArgGlnArgTyrSer 748  
 Db 41419 AGCGGACCGCTTCTCGTGGTGGCTGATGCCAACCTGGGGAGCTCGCAACGCTACTCT 41360  
 QY 749 LeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGly 768  
 Db 41359 ATCACTACTCCAAACACCGCTGGGGTTCAGATGTGCAATTCCTCGCTTTCACGGCGCAG 41300  
 QY 769 IleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMet 788  
 Db 41299 AACGCTGGATCCGACACCTAGCATCGC-----CATCGTTT 41261  
 QY 789 ThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnHisValProTyrArg 808  
 Db 41260 GTTACACGGGACCGCTGGCTGATGAACCGGCTGATTCGACAAATACCCCGCAT 41201  
 QY 809 LeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828  
 Db 41200 CTGGCTTCTTCTCGCGGGGACCGCAGTATTCGTGGCTCAAAATACAA-TCTATCGCT 41142  
 QY 829 ProIleSerAspLysGlyTyrLeuThrGly-----GlyGlnValLeuAlaVal 844  
 Db 41141 CCGAAATACCGCAGCGGTGACTGAAGAGGCTCGAAGTGATACCGGATC----- 41091  
 QY 845 GlyThrAlaGlyTyrAsnTyr 851  
 Db 41090 -----GCTGAATACCACTAC 41076  
 RESULT 35  
 NMA122491 340806 bp DNA linear BCT 04-DEC-2000  
 LOCUS NMA122491  
 DEFINITION Neisseria meningitidis serogroup A strain 22491 complete genome;  
 segment 1/7.  
 ACCESSION AL162752 AL157959  
 VERSION AL162752.2 GI:7378778  
 KEYWORDS  
 ORGANISM  
 SOURCE Neisseria meningitidis 22491.  
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
 Neisseria.  
 REFERENCE  
 1 (bases 1 to 340806)

AUTHORS  
 Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,P., Felwell,T., Hamlin,N., Holtroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,W., Skellon,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.  
 TITLE  
 Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491  
 JOURNAL  
 Nature 404 (6777), 502-506 (2000)  
 MEDLINE  
 20222556  
 PUBMED  
 10761919  
 REFERENCE  
 2 (bases 1 to 340806)  
 Parkhill,J.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk  
 COMMENT  
 Notes:  
 Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web.  
 (URL, [http://www.sanger.ac.uk/Projects/N\\_meningitidis/](http://www.sanger.ac.uk/Projects/N_meningitidis/)).  
 FEATURES  
 source  
 location/Qualifiers  
 1..340806  
 /organism="Neisseria meningitidis 22491"  
 /strain="22491"  
 /db\_xref="taxon:122587"  
 /note="serogroup: A"  
 complement(38..358)  
 /gene="NMA0001"  
 complement(38..358)  
 /note="NMA0001, unknown, len: 106 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein NMA0001"  
 /protein\_id="CAB83521.1"  
 /db\_xref="GI:7378779"  
 /db\_xref="SPTREMBL:Q9JX93"  
 /translation="MSNTQITRSADPTSRAGCALDIANNNGTVALHTDQPYKWHYNDGEVFAVMCEVDVMHYRENEEHHVRLKSDIFPAGTGEHVAHPRGRTLVLEKBCSV"  
 complement(414..2438)  
 /gene="nuoL"  
 complement(414..2438)  
 /gene="nuoL"  
 /BC\_number="1.6.5.3"  
 /note="NMA0002, nuoL, NADH dehydrogenase I chain L, len: 674 aa; similar to many e.g. NUOL\_RHOCA NADH dehydrogenase I chain L (BC 1.6.5.3) (700 aa). Contains Pfam match to entry PF00361 oxidored\_g1, NADH-ubiquinone/plastoquinone (complex I), various chains, and to entry PF00662 oxidored\_g1\_N, NADH-ubiquinone oxidoreductase (complex I), chain 5 N-terminus"  
 /codon\_start=1  
 /transl\_table=11  
 /product="NADH dehydrogenase I chain L"  
 /protein\_id="CAB83322.1"  
 /db\_xref="GI:7378780"  
 /db\_xref="SWISS-PROT:Q9JX92"  
 /translation="WMDITLYLIALVPLAGSLAGLEGKIGRGAHTVTTLGVAVS AVLSAVYLMGFLNGSRKAFDENVTYMLMGGLDFSVGLVDNTAMMYVGVSLAV HIYITGVHNDKGVYQREFSYISLFTFSMLIMSNFOLFPGWEAVGLVYLIGF YKPRSGAHNFAKPLINVGDFGFLGLGLVAVFGSLRYODVAYLPVONNATIO LEPGVMSLITVTCLILPFGAKSGAOPFLHWLPDSMGSPRISALIHAAATMVAGL FVMSRSPYTEMSTALSYIYIGAITLAFKPFLEYIIONDKRVAYSTISLQWITV ALGASAYVAMFHVMTAFKALPLELAAGSALITGHHDDMRKMLKRYMTITWTM LIGNLSLGTPEPFGYSKDSIIEAKYSTLPGSGFAVAVAVFAVFAFROFM VEHGEKRSRLPEHNSDGHSEHNGKNDNHESPLVTLPLILAVSVIIGYAI EPMVLGDFEKDVIYFNADAHPTMKKEFHGALAMVSHSLSPVLVYLAIAGVLSML LYKRLPHAKIAQAFRPYVYLFENKYYIDALFYNFPAKTRALGTFFPKVGDALID NGIVNGSRVLGAVAAOYKRVDTGTFTYTAAMVFGVLLDMTETGLR"  
 1035..1044  
 /note="Core DNA uptake sequence: gccgtctgaa"

misc\_feature

```

misc_feature      /label=DUS
                  complement(1116..2030)
                  /gene="nuoL"
                  /note="Pfam match to entry PF00361 oxidored_g1,
                  NADH-ubiquinone/plastoquinone (complex I), various chains,
                  score 351.80, E-value 7.3e-102"
                  complement(2061..2243)
                  /gene="nuoL"
                  /note="Pfam match to entry PF00662 oxidored_g1_N,
                  NADH-ubiquinone oxidoreductase (complex I), chain 5
                  N-terminus, score 77.20, E-value 3.5e-19"
                  complement(2487..2828)
                  /gene="NMA0003"
                  complement(2487..2828)
                  /gene="NMA0003"
                  /note="NMA0003, unknown, len: 113 aa"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0003"
                  /protein_id="CAB83323.1"
                  /db_xref="GI:7378781"
                  /db_xref="SPTREMBL:Q9JX91"
                  /translation="MMDKNOLEDFHKLMLNLYOEALNLPQYKATRFLOIVNFGCK
                  EAADKLSTGEKKQTGTETELISGGVHALKYSMEYLVLOKPMCDLFTBEOALVARK
                  RLERYGVSPK"
                  complement(2833..2837)
                  complement(2952..3527)
                  /gene="NMA0004"
                  complement(2952..3527)
                  /gene="NMA0004"
                  /note="NMA0004, unknown, len: 191 aa; similar to
                  hypothetical proteins e.g. Y977_HAEIN HI0977 (191 aa),
                  fasta scores: E(): 0: 80.5% identity in 190 aa overlap,
                  and to proteins involved in cell cycle e.g. TR:054679
                  (EMBL:AF036487) Laccococcus lactis plasmid pNZ4000
                  putative mobilization protein (200 aa), fasta scores: E():
                  0, 55.1% identity in 167 aa overlap, and FIC ECO1 cell
                  filamentation protein FIC (200 aa), fasta scores: E():
                  0.093, 29.4% identity in 126 aa overlap"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0004"
                  /protein_id="CAB83324.1"
                  /db_xref="GI:7378782"
                  /db_xref="SPTREMBL:Q9J0R9"
                  /translation="MSENPICKTKMSIDIOSLHNRARFESGDDIRIVGTAGIQO
                  IRRYLGGLYDFAGQIREDNISGGRFANMAYLKALYKIDMPERTFEETIAKVE
                  MNIAHPLEGNRSTRIMDLVKKLVVMQNVSKTLYLOAMERSPVNDLEIRFL
                  LKDNLTDDVNNKEIRKIGIEQSYIEYEGK"
                  complement(3516..3525)
                  /gene="NMA0004"
                  /note="Core DNA uptake sequence: gccgctctgaa"
                  /label=DUS
                  complement(3561..3866)
                  /gene="nuoK"
                  complement(3561..3866)
                  /gene="nuoK"
                  /EC_number="1.6.5.3"
                  /note="NMA0005, nuoK, NADH dehydrogenase I chain K, len:
                  101 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase
                  I chain K (EC 1.6.5.3) (102 aa), fasta scores: E():
                  1.1e+24, 65.0% identity in 100 aa overlap. Contains Pfam
                  match to entry PF00420 oxidored_g2,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 4L"
                  /codon_start=1
                  /transl_table=11
                  /product="NADH dehydrogenase I chain K"
                  /protein_id="CAB83325.1"
                  /db_xref="GI:7378783"
                  /db_xref="SPTREMBL:Q9J0Q9"
                  /translation="MITLTHYLVGALLFGISAMGIFMNRKNVLLMSTIEMLAVN
                  FNFIAFSQHLGDTGQIFVFPVLVVAAESAIGLAIVLVYRNQOTINVADDELKG"
                  complement(3564..3863)

```

```

gene              /gene="nuoK"
                  /note="Pfam match to entry PF00420 oxidored_g2,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 4L,
                  score 141.50, E-value 1.5e-38"
                  complement(3863..4534)
                  /gene="nuoJ"
                  complement(3863..4534)
                  /gene="nuoJ"
                  /EC_number="1.6.5.3"
                  /note="NMA0006, nuoJ, NADH dehydrogenase I chain J, len:
                  223 aa; similar to many e.g. NUOL_RHOCA NADH dehydrogenase
                  I chain J (EC 1.6.5.3) (202 aa), fasta scores: E():
                  9.7e+21, 41.3% identity in 201 aa overlap. Contains Pfam
                  match to entry PF00499 oxidored_g3,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 6"
                  /codon_start=1
                  /transl_table=11
                  /product="NADH dehydrogenase I chain J"
                  /protein_id="CAB83326.1"
                  /db_xref="GI:7378784"
                  /db_xref="SPTREMBL:Q9JX90"
                  /translation="MTFOLIFKIFAVIILYGAIKTVTKPNVNAHLIVTPCVSAM
                  LMLMQAEFGVTLVVYGVAVVLFVYVMMINIDIEBRAGFNHARVAGVGTLL
                  AVALLILVNPRTDLAAGIMKIDIPADYNNIRGLSGRITDILPELAAVLILGMV
                  AATAIVHRKTVNPKRMDPADQVVRADQGRHRLVMEAVKPTQESAEESEVSDLKPK
                  EBGKA"
                  complement(3874..3879)
                  /gene="nuoJ"
                  complement(4034..4525)
                  /gene="nuoJ"
                  /note="Pfam match to entry PF00499 oxidored_g3,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 6,
                  score 77.90, E-value 2e-19"
                  complement(4570..5277)
                  /gene="NMA0007"
                  complement(4570..5277)
                  /gene="NMA0007"
                  /note="NMA0007, unknown, len: 235 aa"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0007"
                  /protein_id="CAB83327.1"
                  /db_xref="GI:7378785"

RBS
                  complement(3874..3879)
                  /gene="nuoJ"
                  complement(4034..4525)
                  /gene="nuoJ"

misc_feature
                  /note="Pfam match to entry PF00499 oxidored_g3,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 6,
                  score 77.90, E-value 2e-19"
                  complement(4570..5277)
                  /gene="NMA0007"
                  complement(4570..5277)
                  /gene="NMA0007"
                  /note="NMA0007, unknown, len: 235 aa"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0007"
                  /protein_id="CAB83327.1"
                  /db_xref="GI:7378785"

gene
                  complement(4570..5277)
                  /gene="NMA0007"
                  complement(4570..5277)
                  /gene="NMA0007"
                  /note="NMA0007, unknown, len: 235 aa"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0007"
                  /protein_id="CAB83327.1"
                  /db_xref="GI:7378785"

CDS
                  complement(4570..5277)
                  /gene="NMA0007"
                  complement(4570..5277)
                  /gene="NMA0007"
                  /note="NMA0007, unknown, len: 235 aa"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0007"
                  /protein_id="CAB83327.1"
                  /db_xref="GI:7378785"

Alignment Scores:
Pred. No.: 0.000159      Length: 340806
Score: 259.00           Matches: 191
Percent Similarity: 33.48%      Conservative: 121
Best Local Similarity: 20.49%    Mismatches: 342
Query Match: 5.48%           Indels: 278
DB: 1                     Gaps: 40

US-09-914-168-2 (1-919) x NMA122491 (1-340806)
QY 116 ThrProLeuSerLeuGluLeuPheAlaGluGlnSerThrGluMetGlyIleAsnPro 135
Db 84027 TCGCCTTGCGACTTGCCGACTTCACATCCAAAGCATCCGCGTGAGGCGTTGACGGT 84086
QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGlnProAsnSer 149
Db 84087 ACCGAGCCGAGTACCGGTATTCACATCACTGCCGCAAGTCCGGGACACCTACAAAC--- 84143
QY 150 GluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 84144 -----GACACACAGCGCAATGCCATCATCAAAAGCTGTGAC 84179
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
Db 84180 GCCACCGGTTCTTGTGACGAGCTACGCGTCAAACTCGCGAGCGGAGCTCTGTCGAC 84239
QY 187 PheTyrGlnSerSerGlnSerGlyGlnThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 84240 GTATGCAACGCCCCACATGAGGCTCGCTCAACATCAACGCGGCAAAATGTCGCAAAAC 84299

```

QY 207 GtuproTyrAlaAsnIleLysAlaIleuGluAsp-----IleThrGlnIleuSerAla 224  
 Db 84300 GAGCCC-----ATTAAACAAAACCTCGAATGCTTGGGGCTGGCCAGTCGCAATAC 84330  
 QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236  
 Db 84351 TTTAATACGGCGACATCAATACAGCATGCGCGGCTGAAAGAAATACCTCGGGCGC 84410  
 QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249  
 Db 84411 GGCAAACTCAATATCAAAATACAGCCCAAGTACCAAACTCGCCGACCGCTCAC 84470  
 QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269  
 Db 84471 ATGCACATCAAGATTGACGAGGCGCAATCCGCCAAATACACGACATGGAATTTCAA--- 84527  
 QY 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289  
 Db 84528 GGCACCAACAGCTGTATTCGCGCAACCTGATGCGG----- 84563  
 QY 290 LysAlaIleThrIleThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306  
 Db 84564 -----CAGATGTCGCTGACCGCAAGGCGGATTTGGACATGG 84599  
 QY 307 -----PheHisIleGlyLysTyrGluThrLysLysAsnLeuIleGlu 320  
 Db 84600 CTGACAGCAAGCAACCAATTCACGACAGCAAAATTTGCCCAAGACATGGAAAAAGTAAAC 84659  
 QY 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340  
 Db 84660 GACTTCTACCGAAGACAGCGCTACTGTTCGCTATTCGATCCGATACCGATCCCAAC 84719  
 QY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360  
 Db 84720 AACCAAGCAAAACCAACAGACATCAAAATACACGCTCCAGCAAGGCGGACGTTTCGT 84779  
 QY 361 PheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380  
 Db 84780 TGGGCAAAAGTC-----TCCATCGAAGGCGACACCAACGAA----- 84815  
 QY 381 AspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGlu 400  
 Db 84816 -----GTCGCCAAAGCCGAACTGGAAAAAATGCTGACCATGACCCGCGCAA 84863  
 QY 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420  
 Db 84864 TGTACGAAAGCCAGACATACCGCGCTTTGGGT----- 84899  
 QY 421 AsnMetValAsnThrGluIleValIlePheProGluArgGlnIleGlnAsnAspGlnVal 440  
 Db 84900 -----GAGATTCAAGACCGCATGGGC 84920  
 QY 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460  
 Db 84921 TCGCGAGGCTACGATACGC-----GAATTCAGCGTA 84953  
 QY 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479  
 Db 84954 CAGCGCGTGGCCAAAGCCGAAACCAAAACCGTCATTCGCTCTGCATCAACCGGCGC 85013  
 QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492  
 Db 85014 CGGAAATCTACGTCAACGAATATCCATCACCGCGCAACAAACCGCGACGAA--- 85070  
 QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508  
 Db 85071 ---GTCGTGCGCGCGCAATTCGCGCAATTCGCGCTTACGACCACTCCAAAGCTG 85127  
 QY 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523  
 Db 85128 CAACGCTCCAAAGAGCGGCTGAGCTTTGGGCTACTTCGACACACGTACAG----- 85178

QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543  
 Db 85179 -----TTGATGCCGTCGCCCTTGGCGGACACCGCCGCAAA--- 85214  
 QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaAlaAsnArgLysThrProAla 563  
 Db 85215 -----GTCGATTTG-----AACATGAGCTGACCGAAGCTTCCACCC----- 85250  
 QY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
 Db 85250 ----- 85250  
 QY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601  
 Db 85251 ---GGCTGCTGACTTGAAGCGCGCGCTGGGTACAGATATACCGC-----CTGGTCATG 85301  
 QY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
 Db 85302 TCCGAGCGGCTTCCAGACCAACCTGTTCGATACGCGCAAGTCGCGCCCTGCGCGCGC 85361  
 QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
 Db 85362 TCACGAGCAAA-----ACCAAGCTCAACGCGCTGCTGTTTACGACCCGCTTTC 85415  
 QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661  
 Db 85416 ACGGACAGCGGGGTACAGCTGGGCTAC-----GATGTTTACGAAACCTTCGACCGG 85469  
 QY 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681  
 Db 85470 CCGAAAGCTACGACGACGATCAAAACATATAAACCCACGACGCA-----GCG 85517  
 QY 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696  
 Db 85518 GCAGGACATCCGATGACGCTGCTGTACCAATACGACCGCGTGAATTTCCGTTTGGTG 85577  
 QY 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713  
 Db 85578 GCAGAAACCTGACCGCTACACCTACACAAAGCCCAACACATATGCCGACTTATTC 85637  
 QY 714 Asn-----GlyLys-----ProSerGlnGluAlaLeuLeuAlaGlyAlaValHisLys 730  
 Db 85638 AAGAAATACGGGCAAAACCGACGACGACGCGCTTCAAAAGCGCTGTACAAAGCT 85697  
 QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetLeuArgGlyTyrArg 744  
 Db 85698 ACCGTGCGCTGGGGCGCGCACAAACCGACAGCGCTTATGGCGGACGCGGCTAC--- 85754  
 QY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764  
 Db 85755 -----CTGACGGCGCTGAACCGCGCAATC 85778  
 QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
 Db 85779 GCCCTGCCCGCGCAGCAAACTGCAATAC-----TACTCCGCC 85814  
 QY 785 AlaHisGlnMetThrGly----- 790  
 Db 85815 ACCCAACACCAAACTGGTTTCCCTTAAGCAAAACCTTACCGCTGATGCTGGCGGC 85874  
 QY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808  
 Db 85875 GAAGTGGGCAATGGCGGGCGGTAC-----GCGAAGAACCAAAAGAAATTCCTCTCTT 85925  
 QY 809 LeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828  
 Db 85926 GAAACCTTCTACGCGCGCGCTGGGCTGCTGCGCGGTATACGAAAGCGGACAGCTCGCT 85985  
 QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGly----- 845  
 Db 85986 CCG-----AAAGTATATGACGATATACGGCGAAAAATACAGCTACGCGGCAACAA 86036  
 QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856

```

Db 86037 AAGGCGAAGCTCTCCGCGAGCTGCTTCCGATGCCCGGCGGCGAAGCGCGGCGAC 86096
Oy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 86097 GTCCGCTGAGCGCTTTTCCGAGCGACGACGCTGTGGACGCAAAACCTACGACGAC 86156
Oy 870 ----- 870
Db 86157 AACACGACTCCGCGACGCGCGGCGAGGTTCAAAATTTACGCGCGCGCAATACCAT 86216
Oy 871 ---LysGlyPheThrAsnAspRhrIleGlyAlaGlyValAlaArgTrrAlaSer 889
Db 86217 AAATCCACTTTACCAAGCAATTTGCGGTATTCGCGCGCGCGGCTTACCTGCTCG 86276
Oy 890 ProValIleGlyValAlaArgValAlaSerValAlaThrGlyValLysGly-----GluGlyAsn 907
Db 86277 CCTTACGCGCGGATGAATTCAGCTACGCTTACCGCTGAGCAAAACCGAAGACGAA 86336
Oy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 86337 ATCCAAACGCTTCCAAATTCGACGCGACGACGCTTC 86372

```

```

RESULT 36
AX043922/c AX043922 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 1 from Patent WO0066791.
DEFINITION AX043922
ACCESSION AX043922.1 GI:11342850
VERSION
KEYWORDS
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

```

```

REFERENCE 1 (bases 1 to 349980)
AUTHORS Pitzia M., Hickey E., Peterson J., Tettelin H., Venter J.C.,
Mastigiani V., Galeotti C., Mora M., Ratti G., Scarselli M.,
Scariello V., Rappunli R., Frazer C.M. and Grandi G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 1 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
source 1. 349980
location/Qualifiers

```

```

/organism="Neisseria meningitidis"
/dxref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases--seq 108: 300001 to 649980 349980
bases--seq 109: 600001 to 949980 349980 bases--seq 110:
900001 to 1249980 349980 bases--seq 111: 120001 to 1549980
349980 bases--seq 112: 1500001 to 1849980 349980 bases--seq
113: 1800001 to 2149980 349980 bases--seq 114: 2100001 to
2273325 172325 bases"

```

```

BASE COUNT 83241 a 85091 c 95206 g 86442 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 0.000164 Length: 349980
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: 6 Gaps: 40

```

```

US-09-914-168-2 (1-919) x AX043922 (1-349980)

```

```

Oy 116 ThrProLeuSerLeuGluLeuPheAlaGluLeuSerThrGluMetGlyIleAsnPro 135
Db 180761 TCGCCTTGCGACTTCCGCGACTTCCACATCCAGACATCCGCGTGAAGCGCTTGCAGCGT 180702
Oy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
Db 180701 ACCGAGCGGACGATACCTCACTACCTGCGCCCTCAAAAGTCGGGACACCTACACAC--- 180645

```

```

Oy 150 GluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 180644 -----GACACACACGCGACATGCGCATTCATCAAAAGCGCTGAC 180609
Oy 170 AlaArgLeuPheAsnAspGlyValAlaAsnLysValPro-----ArgLeuLysAlaLys 186
Db 180608 GCCACCGGCTTCTTGGAGAGACATGCGCGTCAAACTCGCGGACGCGAGCTCTGTGACC 180549
Oy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 180548 GTTATCGAACGCCGCCACCATGGCTCGCTCAACATCAACGCGCGCAAAATCTCTCAAAAC 180489
Oy 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnLysAla 224
Db 180488 GACGCG-----ATTAAAGAAAACCTCGAATGCTTGGCGGCGCGCACTGCCAATAC 180438
Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuAlaGln----- 236
Db 180437 TTTAATCAGCGGACACATCAACGACGACGCTCGCGCGTGAAGAAAGAAATACCTCGGCGC 180378
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
Db 180377 GCGAAACTCATATTCCAATTCACGCGCGCAAAAGTAACCAACCTCGCGCAACCGCGGTGAC 180318
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAlaAspValIleIleHisAspLeu 269
Db 180317 ATCCGACATCAGATTTGACGAGCGCAATCCGCGCAAAATACCGCAATTCGAATTTGAA--- 180261
Oy 270 GlyGluProValTyrIleAspTyrArgValAlaGluValAlaArgGlyGluGlyAlaAspAsp 289
Db 180260 GGCACCAAGTCATTTCCGCGCGCAACACTGATGCGG----- 180225
Oy 290 LysAlaPheThrThrValAlaAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 180224 -----CAATGTCCCTGACCGCAAGCGCGCATTTGACATGCG 180189
Oy 307 -----PheHisIleGlyLysTyrGlyGluThrLysLysAsnLeuIleGlu 320
Db 180188 CTGACACGACGAACCAATCAATCAACGAGCAAAATTTGCCAAGATATGAAAAAGTAACC 180129
Oy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrrPheAsnArgSerValAlaAspVal 340
Db 180128 GACTTTCACCAAAATTAACGCTACTTCGATTTCCGATTCCTCGATACCGACATCCAAAC 180069
Oy 341 IleLeuProAspAsnThrAlaAlaAspValSerLeuIleTyrAspThrGlnTyrArg 360
Db 180068 AACGACGACAAACCAACGACGACATCAAAATCAACGCGCGCAAGCGGACGTTCCGT 180009
Oy 361 PheAspGlyValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 180008 TGGGCGCAAAATGTC-----TCCATCGAAGCGGACACACCAACGAA----- 179973
Oy 381 AspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuPheValAsnMetGlyGlu 400
Db 179972 -----GTCCCAAGCGGACCTGGAATAAATCTGACCATGAAACCGCGCAAA 179925
Oy 401 AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 179924 TGGTACGAAACCCAGAGATGACCGCGCTTTGGGT----- 179889
Oy 421 AsnMetValAsnThrGlnIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
Db 179888 -----GAGATTCGAAACCGCATGCGC 179868
Oy 441 SerPheGlnIleSerSerSerArgThrGluProAlaGlnValAlaAspGluSerThrLeu 460
Db 179867 TCGGACGAGCTACGATACAGC-----GAATACAGCGTA 179835
Oy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
Db 179834 CAGCGCGTGGCGAAGCGGTGAACCAAAACGTCGATTTGCTCGACATCGAAGCGGCG 179775
Oy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492

```



gene RGLCLVLLFMFFALFNDYFCWFS"  
1100..3731  
/gene-"omp89"  
misc\_binding 1100..1117  
/gene-"omp89"  
/note-"putative Fur-box"  
/bound\_molecule-"Fur"  
1335..3731  
/gene-"omp89"  
/note-"hemin-binding ability: similar to Haemophilus  
influenzae bacterial surface antigen AG15"  
/codon\_start=-1  
/transl\_table=-11  
/product-"Omp89"  
/protein\_id-"AAL66374.1"  
/db\_xref="GI:18252649"  
/translation="MTMNSRFLNAAVLVIGMVRVIAPTTAFVSIAMVEEVOASVRSI  
EVHGKPFVSNQVVRNDIDKVKSPSSGGVDPAVKRLFALGFLPYDKIKQVGRLYVL  
VREYVNOVLFQGNKSLKDPDLKRFISLKPENPSAKLSADVNIKREYKTVGRND  
IAVTQITNLGKRVNVFNIVEGRRTKISNTIFKNHAFSGSLRLVDISTKSGILS  
LLRGDVSSEERLADAEALRFYVNRGYADFRVVISKAFDEARNAVEIDFVLDEGV  
RKIKGVQIESLIDIDGIDVQSMKGLKTRSGDVSANIEQSVAILNNKAADSGYAFK  
VEPRGNRDLANHTISILYNIEGSPRAYVORIEIRGNEKTRDYIRREIDLNEGDVNO  
TIVORAKRRLSEIGFFKAVNISMPDOPDQITLVYDVEAPTDGLESFGGYTGGTS  
PGVSLSEVSTERNLGRCGYVRIGLAGOGEKSNYLSYDPTFLGRLSAGYDIFRS  
TRADAIYVRQTSGLRFRAPINDOLSNLAISYIOEYDFGKTKDLSKETDIRELY  
GRYSGAIVQAQHSPPKRSISYGLYNTIDDKNPHDMGVAVLQEVAGLGNARFL  
KTTGKAMTKTSLDRMDLVGLSFGGVTHEIGREGVRFDMKINTDMIKGRKNGI  
GPROVANKGEVFLGCTTYMNATAEVOFPFIPVPEGELRGALFADVATLYGSNYRP  
FOCETPVNTIKSAMRSSAGVSLMWDSPFGRPLRFDVAMPKKKGQDRLOQLNGISIKF  
"

gene 3798..4844  
/gene-"lpxd"  
CDS 3798..4844  
/gene-"lpxd"  
/function="putative involvement in lipid A biosynthesis"  
/codon\_start=-1  
/transl\_table=-11  
/product-"Lpxd"  
/protein\_id-"AAL66375.1"  
/db\_xref="GI:18252650"  
/translation="MADTFPPFRRLTVANVAELGAKILNPEFSNTVJNTLSIES  
ACKGLVFEHRRKFDALYGSAAVAFCTNEIYFKPESMAITVSTPDRDPDQIF  
LPPDSKPKPMFGOREISPHAHIHPSAKIAGDVCIFAGAVICRNVTGSGSLASIAV  
VIGENKRICDCYTIAPRVQYSLIGDKVHLVPACIGQCGFRTIGGASISIEKPOLGR  
VIEEDVEIGANTTIDRGFEEDTIIIGEGKINDLVIANNVGRTGLAASOQJGS  
TSIGDMSQLGSGSYGADHIVIGKYVQIAAGSGVMDIPDGEKWSGSPARPFQMFREY  
AALRSICKVKKERK"  
4846..5313  
/gene-"fabz"  
CDS 4846..5313  
/gene-"fabz"  
/function="putative involvement in fatty acid elongation"  
/codon\_start=-1  
/transl\_table=-11  
/product-"Fabz"  
/protein\_id-"AAL66376.1"  
/db\_xref="GI:18252651"  
/translation="MITTEGTRSLFTLIEKLSTLPHRYPPFLIDRIEIDGEGFAI  
GIKNITINEPHFVGRPAKPMVGVLEAMQATAGAILSLRGNCKNTLVYIMTVDN  
AKFRKRVIPGDDKIHIVRLKKRSGMRFRSCIAEVGVAVAEIAMIENE"  
5337..6161  
/gene-"lpxa"  
CDS 5337..6161  
/gene-"lpxa"  
/function="putative involvement lipid A biosynthesis"  
/codon\_start=-1  
/transl\_table=-11  
/product-"Lpxa"  
/protein\_id-"AAL66377.1"  
/db\_xref="GI:18252652"  
/translation="MSGTKIHPTALVEKGAOLGENVFGPFCIHSEAVIIGDCSLMS  
HIVIMKTTLGADSKVFSHALIGAEPDOKHKHGGYTTLSIGNCKTIREVYTHNRGSDS

CDS  
SVGMTIVGNCQFECYAHIAHDCRGNNTVFANNMVIAHVTGVYDIIGGAANVHO  
VRVGHAFIGVSAIVGDIPIPYGTAVQAKLAGLIINIMKRAKGLERKDIHALRHAVA  
MLEHDSKPKPEREVSIVASFPYASQSVVDVNVNFKKGRKRYCTPKREGORIKENKO"  
6164..>6444  
/note-"ORF6"  
/codon\_start=-1  
/transl\_table=-11  
/product-"unknown"  
/protein\_id-"AAL66378.1"  
/db\_xref="GI:18252653"  
/translation="MSVSGARNFSLSRTAIIINGVLPITTAOIELEKNQNPFLVILR  
DEADALVIREHCELSIVELARLVKTLKAAEICNTYLAGGVRRRLPKO"  
BASE COUNT 1732 a 972 c 1583 g 2157 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1 41e-06 Length: 6444  
Score: 255.50 Matches: 205  
Percent Similarity: 36.90% Conservatve: 140  
Best Local Similarity: 21.93% Mismatches: 350  
Query Match: 5.41% Indels: 241  
DB: 1 Gaps: 52

US-09-914-168-2 (1-919) x AF461795 (1-6444) !

QY 92 GluSerProIle-----SerArgIleGlyGluGlnSerProProIleuGlyLeuAsp 108  
||| |||: ||||| ||||| :||| |||  
DB 1326 CAAGAAGCCATGACTATGAATTCAGGTTTAAATGCACACACTGCTGTTGATGAGG 1385

QY 109 MetSerValIleGluGluThrProIleu-----SerIleuGluIleuPhe 124  
||| ||||| ||||| |||||  
DB 1386 ATGAGGCTTATGTGCCCAACACAGCCTTGTGTATCACTGTTGAGAGAGTGCAC 1445

QY 125 AlaGlnIleuSerThrgluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGly 144  
||| |||: |||: |||: |||: |||: |||  
DB 1446 GCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1502

QY 145 GluGlnProAsnSerGluValValProProThrIleuGluPro--GluLysProGly 163  
:||| |||: |||: |||: |||: |||: |||  
DB 1503 GTTCGACATATATATAGACATTAAGGTTTCCAGTGCTGATGTTGATTTT 1562

QY 164 LeuIleuAsnIleuPheuTyrIle--ArgLeuPheAsnAsp--GlyValAsnLysVal-- 180  
||| |||: |||: |||: |||: |||: |||  
DB 1563 GCGGTTAAGCGCTTTTGGCTTAAGGCTTATTTATGATGTTTAAATAAACACAGTACGCC 1622

QY 181 ProArgLeu-----LysAlaLysPheTyrGlnSer 190  
||| |||  
DB 1623 GATAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1682

QY 191 SerGlnSerGlyIleuThrSerAlaIleGlySer-----SerHisGlnLysThrGluPro 208  
||| |||: |||: |||: |||: |||: |||  
DB 1683 AATTAACATCTTAAAGATCCTGATCTTAACGCTTATTTCTTTAAACCAATGAGCCC 1742

QY 209 TyrAlaAsnIleuLysAlaIleuGluAspIleThrglnIleuSerAlaMetLysLeuAsn 228  
:||| |||  
DB 1743 TTCATTTCTGCCAAG-----CTTCA 1763

QY 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyr 248  
:||| |||: |||: |||: |||: |||: |||  
DB 1764 GCGTATCTTATGATGATCTGTA-----GCTTATTAACCTGTTGCTGTAAT 1811

QY 249 AspIleAsnLeuSerIleIleArgAsnSerIleGlyValAsnLysVal--IleIleHis 267  
||| |||: |||: |||: |||: |||: |||  
DB 1812 GATATCGCTTATGATGATCAACTATTAATCTTGGAAGAGCGCGTAATGCTGCTTT 1871

QY 268 AspLeuGlyIleuProValTyrIleAspTyrArgAlaValAlaValArgGlyIleuGlyAla 287  
:||| |||: |||: |||: |||: |||: |||  
DB 1872 AATATTGTTAAAGTGCAGACAGACAAAGATTAGCAATATCTTTAAGGGAATACAGCT 1931

QY 288 --AspAspLysAlaPheThrThrValAlaAspGluValPro----- 300  
||| |||  
DB 1932 TTTCGAGTACGCGCTTACGTATGATTTCAACAAGCCTTCAGGAATACCTTTCGTTA 1991

Tue May 6 09:37:16 2003

us-09-914-168-2.rge

Page 75

```

301 LeuLeuLeuLeuValPheHisIshsclYstyrgIurHrlystysAsnLeuLeuLeu 320
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1992 TTGTTGAGAGTATGCTTATAGTAGAGAACGCTTGCTGTGATGAAGAGCGCTCGT 2051
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
321 AsnAlaSerAlaGluHisIshsclYstyrgIurHrlystysAsnLeuLeuLeu 2111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2052 CGTTTACTATACCGTGTATGAGATTTTCCGTTATTTCCCTTAAGCAGTTT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
341 IleuProAspAsnThrAlaAspValSerLeuLeuLeuLeuLeuLeuLeuLeu 2171
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2112 GATGAAGCCGCTATGCTTATGATGATGCTTCTGCTGATGAGAGTGGCTGATAA 380
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 PheAspGluValValPhePheThrIleAspProLysThrAsnGluLeuThrAsp 2201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2172 ATTGGTATGCTT-----CAGATGAAAGTATATC 400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
381 AspLysLeuProValLysArgLysLeuLeuLeuLeuLeuLeuLeuLeuLeu 2255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2202 GATGGGATGATGCTT-----CAATCTAAGAAAGAAATCTTAAACCCGCTCGGTGAT 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 AlaTyr-----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuLeuAlaThr 2315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2256 GTTATTAGCGCAAAATATTGACAGCTGCTGCAATTATTACAAATTAAGCTGCTGAT 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418 ArgTyrPheAsnMetValAsnThrGluLeuValPheProGlu-----ArgGluGlnLeu 2366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2316 TCGGATATGCTTTGCTTAAGTTGCA-----CCACGAGCAATGCTGATTTGGCA 452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
437 AsnAspGlnValSer-----PheGluGlnSerSerSerSerArgThrGluPro 2426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2367 AATCATACATATTCATTTGTTGATACATTTGACAAAGCTCTCGGCTTATGCTGACGT 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
453 AlaGlnVal-----AspGluSerThrLeuGluProValIle-----GluThrAlaGluLeu 2486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2427 ATTGAATATACGTGTATGAAAGACAGCGATGATGATGATGCGCCGTGAGATGATCTA 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
470 ThrAspGlyIleLeuMetAspLysSerProLeuLeu----- 2546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2487 AATGAAGCGTATCTTACATACAGCTTATGCTGACGCAAGAAACGCTGCTACAGCT 498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
482 -----PheSerAlaSerAsnLeuLeuLeuAspLysLeuLeuValAlaAlaLys 2588
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2547 TTAGGCTTTTAAAGCCCTTAAT-----ATTTCATAGGTCCACCC 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
499 AlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGly 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2589 -----GATCAACCTGATTCAGATTATCTTGTGTAGAT----- 2621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
519 ValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIle 2681
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2622 GTACTTGAAAGCACACAGCGAGATCTTTCTTTCTGAGGGTATACACGCGGTACA 2681
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 2732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2682 AGTCCCTGCTGCTCAGCTTGAAG-----GTATCTGTTACGAGGCTAAT-----CTTGAGCG 2747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
559 ArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyrValPheValAla 2747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2733 CGTGGCTCAG-----TATGTT----- 2747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPglYserAspThrGlyThrArg 2783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2748 -----CGTTTAGGCTTAGGTGCTGCGACAGAGAAATCCCGT--- 2783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
599 LeuValThrLysPheGluHisAsnLeuLeuAsnArgAspGlyTyrGlnAlaGlyAlaGlu 2828
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2784 -----AATTACAAATTTATCTTTGATGATCTTATTTCTAGAGTTATGCT 2828
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
619 LeuArgLeuSerGluAsp-----LysLysGlyValLysLeuTyrAlaThr 633
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2829 TTATCTGCTGGGTGATATTTTTCGACGACTTACCGCGCTGATTAAGACATGATGTA 2888
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
634 LysPro-----LeuSerHisProLeuAsnAspGlnLeuArgAlaThr 647
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```





Db 1090 ---GTGCTGGCGCGCAATTGGCCCAATGGAAATCCGCCCTTACGACACCTCCAAAGCTG 1146  
 Oy 509 -----ArgValIleuAlaIleasnHisaspGlyValaAsnArgSerIle 523  
 Db 1147 CAACGCTCCAAAGAGCGCGCTTGAGCTTTGGCTACTTCCGACAAAGACAG----- 1197  
 Oy 524 LeuGlyArgIleSeraspAlaValaSerAlaValaIleArgAlaIleuProaspGlySer 543  
 Db 1198 -----TTTGAATGCCGCTCCCGCTTCCGCGACACCCGCAAA--- 1233  
 Oy 544 GluAsnGluValIleaspLeuProGluArgThrAlaLeuAlaAsnArgIleThrProAla 563  
 Db 1234 -----GTGCAATTTG-----AACATGAGCCTGACCGACGCTTCACCC----- 1269  
 Oy 564 AspValThrGlnSerLysLysValProLeuThrValPheValAlaSeraspLysProArg 583  
 Db 1269 ----- 1269  
 Oy 584 AspGlyGlnIleGlyLeu-----GlyThrPglySeraspThrGlyThrArgLeuValThr 601  
 Db 1270 ---GGCTCGCTGCACCTTGACGCGGCGGTGACAGATACCGGC-----CTGCTCATG 1320  
 Oy 602 LysPheGluHisasnLeuIleasnArgaspGlyThrGlnAlaGlyAlaGluLeuArgLeu 621  
 Db 1321 TCCGCAAGCGGCTTCCCAAGACAACTGCTTCGACGGGCAAGTCGGCGCGCTCGCGCC 1380  
 Oy 622 SerGlaAspLysLysGlyValaLysLeuThrAlaThrLysProLeuSerHisProLeuAsn 641  
 Db 1381 TCACGAAGCA-----ACCAGCTCAACGCGCTGCTGCTTTCACGACCGGCTACTTC 1434  
 Oy 642 AspGlnLeuArgAlaThrLeuGlyThrGlnGlnGluValPheGlyHisSerThrAsnGly 661  
 Db 1435 ACGGCAAGCGGCTGACGCTGGGCTAC-----GATGTTACGAAAGAAAGCTTCGACCGC 1488  
 Oy 662 PheaspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681  
 Db 1489 CGCAAGACATCGACACGATCAAAATATAAACCCACCGCGCA-----GGC 1536  
 Oy 682 GlyThrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696  
 Db 1537 GCAGGATCCGATGAGCGCTGTTTACGATACGACCGCGCTGAATTCGTTGGTG 1596  
 Oy 697 LysThrGlnAlaProGluThrThrPglyAsnLeuPro-----ValaspPheVal 713  
 Db 1597 GCAGAACCTGACCTGACACCTACAAACAAAGCGCCCAACACTATGCCGACTTATC 1656  
 Oy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuAlaGlyAlaValHisLys 730  
 Db 1657 AAGAAATACGCAAAACCGACGCGACAGCGAGCTTCAAAAGCTGGCTGTACAAAGCT 1716  
 Oy 731 ThrVal-----AlaaspAsnLeuValAsnPrometArgIleTyrArg 744  
 Db 1717 ACCGTGGCGTGGGGCGCAACAAACCGACAGCGCTTATGCGCGCGGCTAC--- 1773  
 Oy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSeraspAlaAsnMetAlaIle 764  
 Db 1774 -----CTGACGGCGCGGACGCGCAATC 1797  
 Oy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
 Db 1798 GCCCTCCCGCGAGCAAACTGCAATAC-----TACTCGGCC 1833  
 Oy 785 AlaHisGlnMetThrGly----- 790  
 Db 1834 ACCCAACCAAACTGGTCTTCCCTTAAGCAAACTTCACGCTGATCTCGCGGC 1893  
 Oy 791 -----GlyIleGlnAlaGlyTyrIleTyrSeraspAsnPheAsnHisValProTyrArg 808  
 Db 1894 GAAGTCGCGATTCGCGCGGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944  
 Oy 809 LeuArgPheheAlaGlyLysAspGlnSerIleArgGlyTyrAlaHisaspSerLeuSer 828  
 Db 1945 GAAACTTTCACGCGCGCGCTGGTTCGTCGCGGATACGAAAGCGCGACGCTCGGT 2004

Oy 829 ProIleSeraspLysGlyThrLeuThrGlyGlnValLeuAlaValGly----- 845  
 Db 2005 CCG-----AAAGTGTATGACGAATACGGCGGAAAAATACACTACGCGCGCAACAA 2055  
 Oy 846 ThrAlaGluTyrAsnTyrGluPheMet-----Lysasp 856  
 Db 2056 AAGCCAACTCTCCCGCGAGCTGCTCTTCCGATGCTGCGCGCAAGACGCGGAC 2115  
 Oy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870  
 Db 2116 GTCCCGCTGACCTGTTGGCGACGACGAGCGGTGGGACGCAAAACCTACGACGAC 2175  
 Oy 870 ----- 870  
 Db 2176 AACACGACTTCCCGACCGCGCGAGGCTTCAAAACATTACGCGCGCAATACCCAT 2235  
 Oy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyAlaGlyAlaArgThrAlaSer 889  
 Db 2236 AATTCACACTTTACCAAGCAATTCGCTATTCGCGCGCGGCGGTACTGCTCTCG 2295  
 Oy 890 ProValGlyGlnValaArgValaAspValaIleThrGlyValLysGlu-----GluGlyAsn 907  
 Db 2296 CTTTATGGCGCGGATGAATTCAGGTACGCTACCGCTGAAGAAAAACCGCAAGACGAA 2355  
 Oy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 2356 ATCCACGCTTCCATTCACACTGCGCAGACGCTTC 2391  
 RESULT 39  
 AC010535 181233 bp DNA 1linear HTG 26-JAN-2000  
 LOCUS AC010535/c  
 DEFINITION Homo sapiens chromosome 16 clone RP11-273P15, LOW-PASS SEQUENCE  
 SAMPLING  
 AC010535  
 VERSION AC010535.2 GI:6758796  
 KEYWORDS HTGS, PHASE0.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Sequencing of Human Chromosome 16  
 JOURNAL Unpublished  
 REFERENCE  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 COMMENT  
 Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 On Jan 26, 2000 this sequence version replaced gi:5882403.  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: <http://www.jgi.doe.gov>  
 \*\*\*\*\*  
 \* NOTE: This record contains 76 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.  
 \* 1 944: contig of 944 bp in length  
 \* 945 1570: gap of unknown length  
 \* 945 1570: contig of 626 bp in length  
 \* 1571 2381: gap of unknown length  
 \* 1571 2381: contig of 811 bp in length  
 \* 1571 gap of unknown length

\* 2382 2793: contig of 412 bp in length  
\* gap of unknown length  
\* 2794 3510: contig of 717 bp in length  
\* gap of unknown length  
\* 3511 4621: contig of 1111 bp in length  
\* gap of unknown length  
\* 4622 5020: contig of 399 bp in length  
\* gap of unknown length  
\* 5021 5969: contig of 949 bp in length  
\* gap of unknown length  
\* 5970 7386: contig of 1417 bp in length  
\* gap of unknown length  
\* 7387 8141: contig of 755 bp in length  
\* gap of unknown length  
\* 8142 9359: contig of 1218 bp in length  
\* gap of unknown length  
\* 9360 10926: contig of 1567 bp in length  
\* gap of unknown length  
\* 10927 11815: contig of 889 bp in length  
\* gap of unknown length  
\* 11816 11979: contig of 164 bp in length  
\* gap of unknown length  
\* 11980 14676: contig of 2697 bp in length  
\* gap of unknown length  
\* 14677 15050: contig of 374 bp in length  
\* gap of unknown length  
\* 15051 16346: contig of 1296 bp in length  
\* gap of unknown length  
\* 16347 16482: contig of 136 bp in length  
\* gap of unknown length  
\* 16483 17522: contig of 1040 bp in length  
\* gap of unknown length  
\* 17523 18676: contig of 1154 bp in length  
\* gap of unknown length  
\* 18677 18784: contig of 108 bp in length  
\* gap of unknown length  
\* 18785 19464: contig of 680 bp in length  
\* gap of unknown length  
\* 19465 20255: contig of 791 bp in length  
\* gap of unknown length  
\* 20256 20329: contig of 74 bp in length  
\* gap of unknown length  
\* 20330 21427: contig of 1098 bp in length  
\* gap of unknown length  
\* 21428 22233: contig of 806 bp in length  
\* gap of unknown length  
\* 22234 22614: contig of 381 bp in length  
\* gap of unknown length  
\* 22615 22693: contig of 79 bp in length  
\* gap of unknown length  
\* 22694 23354: contig of 661 bp in length  
\* gap of unknown length  
\* 23355 24389: contig of 1035 bp in length  
\* gap of unknown length  
\* 24390 25438: contig of 1049 bp in length  
\* gap of unknown length  
\* 25439 26033: contig of 595 bp in length  
\* gap of unknown length  
\* 26034 27034: contig of 1001 bp in length  
\* gap of unknown length  
\* 27035 28286: contig of 1252 bp in length  
\* gap of unknown length  
\* 28287 28664: contig of 378 bp in length  
\* gap of unknown length  
\* 28665 29673: contig of 1009 bp in length  
\* gap of unknown length  
\* 29674 30889: contig of 1216 bp in length  
\* gap of unknown length  
\* 30890 31997: contig of 1108 bp in length  
\* gap of unknown length  
\* 31998 33208: contig of 1211 bp in length  
\* gap of unknown length  
\* 33209 34186: contig of 978 bp in length

\* 34187 36171: gap of unknown length  
\* contig of 1985 bp in length  
\* 36172 37717: gap of unknown length  
\* contig of 1546 bp in length  
\* 37718 39342: gap of unknown length  
\* contig of 1625 bp in length  
\* 39343 41077: gap of unknown length  
\* contig of 1735 bp in length  
\* 41078 42930: gap of unknown length  
\* contig of 1853 bp in length  
\* 42931 45066: gap of unknown length  
\* contig of 2136 bp in length  
\* 45067 46516: gap of unknown length  
\* contig of 1450 bp in length  
\* 46517 48277: gap of unknown length  
\* contig of 1761 bp in length  
\* 48278 50881: gap of unknown length  
\* contig of 2604 bp in length  
\* 50882 52832: gap of unknown length  
\* contig of 1951 bp in length  
\* 52833 54941: gap of unknown length  
\* contig of 2109 bp in length  
\* 54942 56865: gap of unknown length  
\* contig of 1924 bp in length  
\* 56866 59370: gap of unknown length  
\* contig of 2505 bp in length  
\* 59371 62658: gap of unknown length  
\* contig of 3288 bp in length  
\* 62659 64918: gap of unknown length  
\* contig of 2260 bp in length  
\* 64919 67208: gap of unknown length  
\* contig of 2290 bp in length  
\* 67209 69548: gap of unknown length  
\* contig of 2340 bp in length  
\* 69549 72168: gap of unknown length  
\* contig of 2620 bp in length  
\* 72169 74419: gap of unknown length  
\* contig of 2251 bp in length  
\* 74420 78177: gap of unknown length  
\* contig of 3758 bp in length  
\* 78178 81052: gap of unknown length  
\* contig of 2875 bp in length  
\* 81053 85165: gap of unknown length  
\* contig of 4113 bp in length  
\* 85166 90874: gap of unknown length  
\* contig of 5709 bp in length  
\* 90875 94813: gap of unknown length  
\* contig of 3939 bp in length  
\* 94814 98166: gap of unknown length  
\* contig of 3353 bp in length  
\* 98167 102653: gap of unknown length  
\* contig of 4487 bp in length  
\* 102654 108499: gap of unknown length  
\* contig of 5846 bp in length  
\* 108500 113109: gap of unknown length  
\* contig of 4610 bp in length  
\* 113110 118023: gap of unknown length  
\* contig of 4914 bp in length  
\* 118024 126141: gap of unknown length  
\* contig of 8118 bp in length  
\* 126142 134123: gap of unknown length  
\* contig of 7982 bp in length  
\* 134124 141203: gap of unknown length  
\* contig of 7080 bp in length  
\* 141204 149918: gap of unknown length  
\* contig of 8715 bp in length  
\* 149919 158297: gap of unknown length  
\* contig of 8379 bp in length  
\* 158298 167935: gap of unknown length  
\* contig of 9638 bp in length  
\* 167936 181233: gap of unknown length  
\* contig of 13298 bp in length.  
! Location/Qualifiers

FEATURES

Tue May 6 09:37:16 2003

```

source
1. 181233
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-273P15"

BASE COUNT 45888 a 45036 c 44713 g 45408 t 188 others
ORIGIN

Alignment Scores:
Pred. No.: 0.00049 Length: 181233
Score: 245.50 Matches: 130
Percent Similarity: 37.00% Mismatches: 55
Best Local Similarity: 26.00% Indels: 180
Query Match: 5.19% Gaps: 136
DB: 2 Gaps: 21

US-09-914-168-2 (1-919) x AC010535 (1-181233)

OY 501 H1sleuTyfaspmetProaspRarqValleuAlaIleasnHisaspasp----- 517
|||||
Db 26074 CATCTTTAAGCAGGAGCACTGATAGCATCTGACACAAAAGCCATCATATTTCTGGGCC 26015
OY 518 -----GlyValAsnaRserIleleuglyArqIleSerAspAlaValSerAlaValAla 535
|||||
Db 26014 ACATGCTACCCATCTGTGGCACTGCTGGAGATTCTGCGAGTCACATGCACTGCA 25955
OY 536 Arg-----AlaIleuProasp 541
|||||
Db 25954 AATATCATTCATTCCAGGAAGTCACAGTGGATGACCTATCCACCACTTTCTCCACGA 25895
OY 542 GluSerGluAsnGluValIleAspLeuProGluArqThraIleuAlaAsnaRglYstr 561
|||||
Db 25894 GAGCAG-----CATCTCTCCTGACCAAGATGCTCTCCCAAACTCAC 25850
OY 562 ProAlaAspValTyGlnSer-----LysLysValProLeuTyValPheValAlaSer 579
|||||
Db 25849 AATATTAAGCAATTTGGGCACTGTCAGAGATTCTCTTCAGCTGTATGTGAGATCC 25790
OY 580 AspLysProArqAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThraRglu 599
Db 25789 CAAATTACTAGCAGCAGATGGCCCACTGGTGTGG-----CAGACCACTAGACAGCGC 25736
OY 600 ValThrLysPheGlnHisAsnLeuIleAsnaRAspGlyTyGlnAlaGlyAlaGluLeu 619
Db 25735 -----CCACCGGCTATTCGTGCG----- 25718
OY 620 ArgLeuSerGluAspLysLysGlyValLysLeuTyAlaThraLysProLeuSerHisPro 639
Db 25717 -----AGCCAGCCCGAGGTGGCGCAGTACAGAAATCTGCTCTC----- 25679
OY 640 LeuAsnaRsgLInLeuArqAlaThraThrLeuGlyTyTrpGlnGlnValPheGlyHisSerThr 659
Db 25678 -----CTTCCTTCAACCTGGGCAAGAGAGAGAAC-----CATGCTCTC 25637
OY 660 AsnGlyPheAspLeuSerThraRgThrLeuGlnHisGluIleSer----- 674
Db 25636 AATGAATACTCACTGTGGCAGTACATTCTGAGACCAACCCATGAGTACTATTCTCT 25577
OY 674 ----- 674
Db 25576 CTTCGAAGCTACTTTAGCAAAATCTGCTGCTCCATTTAGAGCTGGTGGAGATTTA 25517
OY 675 -----ArgSerIleIle-----GlnAsnGlyGly 682
Db 25516 AAGCTTCAGGAACCCAGAGCTTTATTAATAAGAAATGTAAGTCTGAGCAATTTTGGG 25457
OY 683 TrpAsnaRgThrTySerLeuArqGlyArqLeuAspLysLeuLysThrGlnAlaProPro 702
Db 25456 -----GTGTGACACCGTAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25406
OY 703 GluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnIleValLeu 722
Db 25405 GTCCCGTGG-----CTTCGCCCGGTGNTGTTGTGAGTGAAGAGAGAAAGAGG----- 25358

```

```

OY 723 LeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGly 742
Db 25357 -----ACAGAAAAAACTTAACAGCTGCGCGTGAATAAACCGATTTTTCG- 25311
OY 743 TyrArqGlnArqTySerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsn-Me 762
Db 25310 TTAGCCGCGCGC-----TCCCTTGTTCGCGTTCT-----ATGAGTCGACCTGATATTC 25260
OY 762 ValAlaIleAlaRgAlaGly-----IleSerGlyValTySerPheGlyAspAsnaIaTyR- 780
Db 25259 GGGAGGAGTACGCCCTCTTTTGTGTGGGCAATACAGTGGGCTCCACCAACTATTTGA 25200
OY 781 -----GlySerAsnaRgAlaHisGlnMetThrGlyGlyIleGln----- 793
Db 25199 GACACCCACCGGGGGTTAAACCGATTTCCCGTTTTCACAGCGCAGAACCTGATTC 25140
OY 794 -----Ala-GlyTyTrIleTrpSera 800
Db 25139 GCACACTGTACAGATCGCATGCTTTTGTACACGGCGACGCTGGCTGATTTGAACCG 25080
OY 800 spAsnRheAsnHisValProTyTrArqLeuArqPhePheAlaGlyLysPglInsrtIleA 820
Db 25079 GTGATTTGACAAAGTACCGCGGATCTGCTGCTGCTGCGGGCGCAGCAGTATTC 25020
OY 820 rGgLyTyRAlaHisAspLeuSerProIleSerAspLysGlyTyRLeuThrGlyGlyG 840
Db 25019 GTGCTTACAATATCAAAATCTATCGCTCCGAATACGCCACACGCTACTGAAAGGGCCT 24960
OY 840 InValLeuAlaValAlaGlyThraIleGlyTyAsnTyGlyPheMetLysAspLeuAla 860
Db 24959 CGAATTTGATTAACCGGATCGCTGGAATACAGTACACGTCAGCGGAATATGTCGGGCG 24900
OY 860 lAvalPheGlyAspIleGlyAsnaIaTyRAspLysGlyPheThrAsnSprThrLysIleG 880
Db 24899 CGGTGTTGTGCATATGTCGCGAAGCGGTAAAGCATATTCGCCGCGCAGCATTTAAACCG 24840
OY 880 lYAlaGlyValAlaGlyValArqTrpAlaSerProValGlyValAlaRgValAlaValAT 900
Db 24839 GTACCGGGGTGCGCGCTGCGCTGGAATCGCGGTGCGGCGCAATCAATCATTTGGCG 24780
OY 900 hrGlyVal-----LysGluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
Db 24779 TACCGGTGCGCGATTAAGACGACACGCG-----TTACAGTTTTAACTGCT 24733

RESULT 40
AE012236/c
LOCUS 10029 bp DNA linear BCT 23-MAY-2002
DEFINITION Xanthomonas campestris pv. campestris str. ATCC 33913, section 144
of 460 of the complete genome.
ACCESSION AE012236 AE008922
VERSION AE012236.1 GI:21112424
KEYWORDS
SOURCE
ORGANISM
Xanthomonas campestris pv. campestris str. ATCC 33913.
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xanthomonas.
REFERENCE
1 (bases 1 to 10029)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,R.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Camargo,L.E.A., Camarotte,G., Camarvan,F., Cardozo,J.,
Chambergo,F., Clapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,Y.K., Oliveira,M.C., Oliveira,V.R., Pereira,J.H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezsa,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and

```

TITLE	Kitaajima, J. P.
JOURNAL	Comparison of the genomes of two <i>Xanthomonas</i> pathogens with differing host specificities
MEDLINE	Nature 417 (6887), 459-463 (2002)
PUBMED	12024215
REFERENCE	2 (bases 1 to 10029)
AUTHORS	da Silva, A. C. R., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Ouaguelo, R. B., Monteiro-Vitorello, C. B., Van Sluys, M. A., Almeida Jr., N. F., Alves, L. M. C., do Amaral, A. M., Bertolini, M. C., Chambergo, L. E. A., Camarotte, G., Cannavan, F., Cardoso, J., Chambergro, F., Chapina, L. P., Ciccarelli, R. M. B., Coutinho, L. L., Cursino-Santos, J. R., El-Dorri, H., Farla, J. B., Ferreira, A. J. S., Ferreira, R. C. C., Ferro, M. I. T., Formighieri, E. F., Franco, M. C., Greggio, C. C., Gruber, A., Katsuyama, A. M., Kishi, L. T., Leite Jr., R. P., Lemos, E. G. M., Lemos, M. V. F., Locati, E. C., Machado, M. A., Medeira, A. M. B. N., Martinez-Rossi, N. M., Martins, E. C., Meidanis, J., Menck, C. F. M., Miyaki, C. Y., Moon, D. H., Moreira, L. M., Novo, M. T. M., Okura, V. K., Oliveira, M. C., Oliveira, V. R., Pereira Jr., H. A., Rossi, A., Sena, J. A. D., Silva, C., de Souza, R. F., Spínola, L. A. F., Takita, M. A., Tamura, R. E., Teixeira, E. C., Tezza, R. I. D., Trindade dos Santos, M., Truffi, D., Tsai, S. M., White, F. F., Setubal, J. C. and Kitaajima, J. P.
TITLE	Direct Submission
JOURNAL	Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900, Brazil
FEATURES	
source	1. 10029 Location/Qualifiers /organism="Xanthomonas campestris pv. campestris str. ATCC 33913" /strain="ATCC 33913" /db_xref="ATCC:33913" /db_xref="taxon:190485" /note="pathovar: campestris" complement(72..2528) /gene="oma" /note="XCC1365" complement(72..2528) /gene="oma" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark" /codon_start=1 /transl_table=11 /product="outer membrane antigen" /protein_id="AA040663.1" /db_xref="GI:21112425" /translation="MKLHSEFLSSSYAGTQOHPSTPGCDMTRLPTRRLALALA AGSLPAALAAEPVADIRVADGLORIASGVEFTLPPVNRGDTVDGKVDGSIATL YRGFEDVOLDRQGNILVTYKRAPIINKLTITGKNDIKSEELKGLGIGTEGCT FDRLSLDRVQLQELRTOINNRKATNVEITPTVPSPLDKRNVALAIKGRKAKTRHVL IGRKEFANEDILENWESKEHNMAWYRADOYSKEKMSGLEKLNWYLDKGVDFSV DSTOVSISPGRDMFVSAGVTEGEVOYKISEIKYGTGDTLPQADIERLIVKADIDFSR ALLEYTSDAITNTLSNNGYAFKVNPIPTPNREDRTVAANLOVPGPVSVRRILKVG NSRTSDEVLRERKOFENTWYSOAATIDRSKIRLORLGYSPSVDEMPAVGSDNDVY VYVVKETSSGFVFGLOFSQAFGVTTSVQLSQNNPLGGGRVAVEASRSTFQQRKAFS YTNPFPTDGVSLGYNLSKRELYDSNTAFOFSTNGSAOVVFGVPTTMDTVSAMIG VDSNQITTFPGTTPQALVDYIOAIGDTFAVTRTEGMARVANDRNDPMPRGVORQG LETTLPGSTVEYKLANLOISNWPPIPAIVLNTREFLGSDSGKSSGVITAGDGT RPVTATGLPEYENYAGCTNSVGFEDNTLGPSEAIINGENRGOPLGGSLKTVGSVM YPKLFDSPASRIASFDPFGNYSVDVAFKANELRASTGVALLMRAVPVGISISYAFP LKKEENDEIRLOPFTGCGF"
gene	complement(2525..3871) /gene="XCC1366" complement(2525..3871) /gene="XCC1366" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AA040664.1" /db_xref="GI:21112426"
gene	/translation="MGDFIGSVMMIVISGLVLTVEFHFGHFVWARRCVYKLVRSVGF GKPLMRDRHGTGEFAIAAIPLOGYVKMLDERGEVHPAREQAEFRKTVYORAIIVA ACPIANILLCMVLMAMFVIGKODYSATYGRADGLAAAGLVGEISVITDGSYSVM SDA5MOLTTAAAMRDROYLTAAEEGNSHHTIRLSOLPAGDERVATLQAGMOPM LOPPYVDKYVASAADGVLPGRIVADIGOPFRSAGVPAQLOALGCTGGCMTEYA REDDRLALEIAPRKSPGEGMMGLGVFAATAAPYDSROOYGAFAAPAARIREGKMTA DSIIGMKRMILTGQASVKNISGPVTIARAANASERVDWFLVFLGLSLSLAIINMP IPILDGHLIYVLIIELVKSGSPSERAMIGOVGLAVLAGLAFVNDILGLVPR" complement(3898..5088) /gene="dxr" /note="XCC1367" complement(3898..5088) /gene="dxr" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/transl_table=11 /product="1-deoxy-D-xylulose 5-phosphate reductoisomerase" /protein_id="AA040665.1" /db_xref="GI:21112427" /translation="MSGSSLRVAVFANGSGISASLDVIARHHPERLRAVLSAGSKV DALLALCVLHRPAHVIADAPALVPALRDLRAAGLTTOAHAGDQALDALASDAPCTV VAIIVGAGLSSTLAAAGAKRLILANKRESIVAGELTTRTAAAGAEITIPDSHSA IPOCLASCDMSKGVRRVYIITASGCPGPRCORADLVTPQAQVAHPKMSKRIYSDS ATLMNGLVETIENHLFLGPEQIDVLVHPQSLVSLVFEVDSGLAQGLDPMRTTL AVGLAMPREVESGVGIDLQGRIDFEAPDEAFPCRLADLADLARGSTAAIILMAA NEVAVSAFLQGVKGFALIPALVEHTLTTRQVADTILNLALFADAEARRTRERALNH SLHA" complement(5091..5918) /gene="cdsa" /note="XCC1368" complement(5091..5918) /gene="cdsa" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
gene	/transl_table=11 /product="phosphatidate cytidyllyltransferase" /protein_id="AA040666.1" /db_xref="GI:21112428" /translation="MTQFRIYALIMAPLCAICVLLPTOMLALALILFLTGWEML KLSGIDSLPRVTLMLNLILWIMYMAVSGVSPFOIAIVAGVWGLVALILMRFEN FGDHNGCORAKLAKGLTVAIPNAALVLIHANPKNNMLTLITPYWNAADSAY FGRGRGKHLAPRISPNTIGSLGLVAGIIVACAFEMGLTILAHVPLVLAAY AVLASVGLDFESLILKRNHAGAKSGTIVIGHGVLDRIDGVLAALPVAFGLKEIIGF" complement(5915..6691) /gene="upps" /note="XCC1369" complement(5915..6691) /gene="upps" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
gene	/transl_table=11 /product="undecaprenyl pyrophosphate synthetase" /protein_id="AA040667.1" /db_xref="GI:21112429" /translation="MASTAMHSEPSAAVPRHLAIMDNGRMAORRRRRPVYIGHRAG ARVAVNTIFEDCLDFKGVSAITLFAFSSENMGRPODEVDAIMKLFILHALDGVGELDRG VOVRPFGDSRFEAPLRDMAGAEERTIANTRIVLSIASYGGRODIAAATAALAVEY AAGRILOPDIIDEALLASRYALADLPAPDLFIRPGGTGTRISNGLIMQALTELMFTETL WPEFDAGVLDQALDDYTAGRERRGTLTSAOIAEKATEASSA" complement(6694..7251) /gene="ftr" /note="XCC1370" complement(6694..7251) /gene="ftr" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/codon_start=1 /transl_table=11 /product="ribosome recycling factor" /protein_id="AA040668.1"





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - nucleic search, using frame\_plus\_p2n model

Run on: May 5, 2003, 23:48:35 ; Search time 325 Seconds

(without alignments)  
6367.958 Million cell updates/sec

Title: US-09-914-168-2

Perfect score: 4727

Sequence: 1 MSKPVLFANRSEMPVALAAY.....TGKVEGEPNKLHFFICTPF 919

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlh  
-Q/cgr2.1/USFTO.spool/US09914168/runat\_28042003\_151614\_6775/app\_query.fasta.1.1095  
-DB=N.geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09914168 @CGN.1.1.205 @runat\_28042003\_151614\_6775 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARR.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEX=7

Database :

N.Geneseq.101002: \*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4727	100.0	2760	21	AAA50536	Moraxella catarrha
2	4727	100.0	89047	22	AAF28547	Genomic fragment #
3	4581	96.9	2670	21	AAA50537	Moraxella catarrha
4	604.5	12.8	1374	22	AAF61044	P. putida KT2440-a
5	436.5	9.2	1737	21	AAA50269	Haemophilus influe
6	427.5	9.0	1731	21	AAA50270	Haemophilus influe
7	320	6.8	2193	23	AA593089	DNA encoding novel
8	319	6.7	944	24	ABK5155	Haemophilus influe
9	265.5	5.6	2379	22	AA507278	Neisseria meningit
10	265.5	5.6	2379	22	AAH42129	Neisseria meningit
11	259.5	5.5	2505	21	AAAI1535	DNA encoding sequenc
12	259	5.5	2394	22	AA507277	Neisseria meningit
13	259	5.5	2394	22	AA507279	Neisseria meningit
14	259	5.5	2394	22	AAH42128	Neisseria gonorrhoe
15	259	5.5	2394	22	AAH42130	Nucleotide sequenc
16	259	5.5	52253	21	AAAB1478	N. meningitidis pa
17	259	5.5	349980	21	AAEF21544	Neisseria meningit
18	253	5.4	2394	21	AAAI15156	DNA encoding outer
19	238.5	5.0	6617	22	AAI97964	Lawsonia intracell
20	235.5	5.0	1425	23	AA581897	DNA encoding novel
21	235.5	5.0	1425	23	AA593087	DNA encoding novel
22	229	4.8	2325	20	AAAX91724	Porphyromonas ging
23	229	4.8	2409	20	AAAX91597	Porphyromonas ging
24	224.5	4.7	1830	21	AAAB4507	Neisseria meningit
25	224.5	4.7	1830	21	AAAB4508	Neisseria meningit
26	222.5	4.7	1848	21	AA53624	Neisseria meningit
27	216.5	4.6	1764	21	AAAB4509	Neisseria meningit
28	216.5	4.6	1848	21	AA53623	Neisseria meningit
29	216.5	4.6	92934	21	AAAB1473	N. meningitidis pa
30	216.5	4.6	173325	21	AAAF21613	Neisseria meningit
31	214.5	4.5	2388	24	ABQ72980	Methylococcus caps
32	214.5	4.5	2351	22	AAQ90024	H. capsulatus caps
33	206.5	4.4	2751	24	AA52588	H. pylori Hps120 e
34	205.5	4.3	3012	19	AAV52012	Helicobacter polyp
35	200	4.2	2778	18	AAAB25138	H. pylori cell env
36	200	4.2	2778	20	AAAX75802	H. pylori outer me
37	184.5	3.9	2989	15	AAQ66202	H. influenzae PAK
38	182	3.9	2342	23	AA593745	DNA encoding novel
39	181	3.8	1916	23	AA593088	DNA encoding novel
40	178.5	3.8	2974	15	AAQ66201	H. influenzae SB33
41	176.5	3.7	2442	21	AAZ28550	M. catarrhalis (AT
42	176.5	3.7	2442	21	AAZ28551	M. catarrhalis (AT
43	176.5	3.7	99629	22	AAAF28550	Genomic fragment #
44	174.5	3.7	2019	22	AAAF61065	P. putida KT2440-a
45	173.5	3.7	8922	23	ABL04053	Drosophila melanog

## ALIGNMENTS

RESULT 1

AAA50536 standard: DNA; 2760 BP.

AAA50536;

05-DEC-2000 (first entry)

Moraxella catarrhalis BASB081 gene coding region.

BASB081 gene; infection; otitis media; sinusitis;

KW inflammation; therapy; antibacterial; antiinflammatory; vaccine;

KW diagnosis; ds.

OS Moraxella catarrhalis.

XX Key Location/Qualifiers

XX CDS 1..2670

FT

[illegible]



QY 421 AsnMetValAsnThrGluIleValPheProGluIleArgGluGlnIleGlnAsnAspGlnVal 440  
 |||||  
 Db 1261 AATATGCTGAATACCGAGATGCTTTCCACAGCGGTGAACAGATCCAAAGACCAACTG 1320  
 QY 441 SerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrIleu 460  
 |||||  
 Db 1321 AGCTTTAGAGCATCTTCAAGTAGCCGTAACACAGCAGCAAGTTGATGAAGACACACTT 1380  
 QY 461 GluProValIleGluThrValGluIleThrAspGlyIleLeuMetAspIleSerProIle 480  
 |||||  
 Db 1381 GAACCTGTCATTTGAACCGTTGAGCTACGAGATGGGATTTAATGATTTTCGCCCATC 1440  
 QY 481 GluPheSerAlaSerAsnIleuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg 500  
 |||||  
 Db 1441 GAATTTAGTGATCTAATCTGATTTCAAGCAAGCAATTTGGTGGCTGCCAAGGCTCGC 1500  
 QY 501 HisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
 |||||  
 Db 1501 CATTTAATATGACATGCCGATGATAGGCTGCTGCCATCAATCATGATGATGCGCTAAT 1560  
 QY 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540  
 |||||  
 Db 1561 CGCTCTATTTGGCGAGATACGCGATCGCTATCTGCCGTTGCACGTGCTAATTTACT 1620  
 QY 541 AspGluSerGluAsnGluValIleAspLeuProGluIleArgThrAlaLeuAlaAsnArgLys 560  
 |||||  
 Db 1621 GATCAATCTGAATAATGAGGTATAGATTGGCCGCGCTACCGCATTCGCTAATCGCAAG 1680  
 QY 561 ThrProAlaAspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAsp 580  
 |||||  
 Db 1681 ACCCTCTGCTATGCTATCAAAAGTAAAAAGTGGCGCTATGCTTGTGGCGGTGAT 1740  
 QY 581 LysProArgAspGlyGlnIleGlyLeuGlyTyrPglSerAspThrGlyThrArgLeuVal 600  
 |||||  
 Db 1741 AAACACAGAGATGTCNAATTTGGTTGGCTGGGATGGACACAGGTACCCGCTACTC 1800  
 QY 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620  
 |||||  
 Db 1801 ACAAAATTTGAGCATAAATTTGATTAATCGTATGCGTATCAAGCAGGCGGTAGACTAGA 1860  
 QY 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640  
 |||||  
 Db 1861 CTGCTGAGGATAAAAAGGGGTCAAGTATATGCCACCAACCGCTTAGCCACCTCTTA 1920  
 QY 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsn 660  
 |||||  
 Db 1921 AATGATCAGCTAAAGCAACTTTGGTTATCAACAAGAGTTTGGTCACTTATCCAAAT 1980  
 QY 661 GlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsn 680  
 |||||  
 Db 1981 GGTTTGATTTATCCACACGACCCCTAGAGCATGATTAACCGCAGATATATCCAAAT 2040  
 QY 681 GlyGlyTrrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAla 700  
 |||||  
 Db 2041 GGTGGCTGGAAATCTACTTATTCATTTGCTTATCGTCTTGATTAAGCTTAAAAACCAACA 2100  
 QY 701 ProProGluThrTrrPglAspLeuProValAspPheValAsnGlyLysProSerGlnGlu 720  
 |||||  
 Db 2101 CCCCCCTGAACATGCGAGATTTTACACGTGATTTTGTCAATGTGTAGCCCAAGCCAAAG 2160  
 QY 721 AlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet 740  
 |||||  
 Db 2161 GCGTTATTGGCAGGTGTGCTGCTCATTAACGTTGAGATTAATTTGTTATATCCGATG 2220  
 QY 741 ArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAla 760  
 |||||  
 Db 2221 CGTGGCATATGTCAGCATATTTCTTATAGAGTTGGCTCAACCGGTTTGGTATCGATCTT 2280  
 QY 761 AsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 780  
 |||||  
 Db 2281 AATATGCTATTTGCTCGAGCTGCTATTAAGTGGCTGATTAATTTTGGGAGTAATGCTTAT 2340

QY 791 GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrrSerAsp 800  
 |||||  
 Db 2341 GGCACCAATCGTGGCCATACAGATGAGTGGTGGCATCAACAAGCAGATATTGGTGGAT 2400  
 QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyValAspGlnSerIleArg 820  
 |||||  
 Db 2401 AATTTTAATCATGTCCATATCGTTGGTTTTTTTGGCTGGCGACCAAGATTTGCT 2460  
 QY 821 GlyTrrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGln 840  
 |||||  
 Db 2461 GCATATGCACATGACAGATTTATACCTATATCAGATTAAGGTTATCTGACAGCGCTCA 2520  
 QY 841 ValLeuAlaValGlyTrrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860  
 |||||  
 Db 2521 GTATTGGCGGTGGTACAGCTCAATATATTTGATTAATTAAGATTTGGCTTTGGCG 2580  
 QY 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880  
 |||||  
 Db 2581 GTTTTGGTGAATTTGGTAATGCTTATGATAAAGCCTTACTAATGATACCAAAATTTGCT 2640  
 QY 881 AlaGlyValGlyValArgTrrAlaSerProValGlyGlnValArgValAspValAlaThr 900  
 |||||  
 Db 2641 GCAGGTGCGGTGCTTCCGCTGGGCAATCACCCTGTCGGTCAAGTTGCTGATGATGCGCACT 2700  
 QY 901 GlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 |||||  
 Db 2701 GGTGTCAAGAAGAGCGCAATCCATTAAGCTGATTTTATTTGGCACACCACTT 2757  
 RESULT 2  
 AAF28547  
 ID AAF28547 standard; DNA: 89047 BP.  
 AC AAF28547;  
 XX 04-APR-2001 (first entry)  
 DT Genomic fragment #34.  
 DE Genomic library: bacteria; human upper airway; otitis media; sinusitis;  
 KW bronchopulmonary; endocarditis; meningitis; ss.  
 XX Moraxella catarrhalis.  
 OS WO200078968-A2.  
 PN 28-DEC-2000.  
 PD 16-JUN-2000; 2000WO-US16649.  
 PP 18-JUN-1999; 99US-0140121.  
 PR (INCY-) INCYTE GENOMICS INC.  
 PA Lagace RE, Patterson C, Berg KL;  
 PI WPI: 2001-041427/05.  
 DR  
 XX  
 PT Genomic library for identifying diagnostic and therapeutic  
 PT compositions, and for identifying virulence factors, regulatory  
 PT elements and drug targets, comprises Moraxella catarrhalis nucleic  
 PT acids -  
 XX  
 PS Claim 1; Page 324-345; 545pp; English.  
 XX  
 CC The present invention relates to a Moraxella catarrhalis genomic library  
 CC comprising of a combination of 41 nucleic acid molecules (see  
 CC AAF28514-AAF28554). The library has a number of uses described in the  
 CC specification e.g. is useful for identifying diagnostic and therapeutic  
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large  
 CC aerobic, gram-negative diplococcus, normally found among the bacterial  
 CC flora of human upper airways. M. catarrhalis is known to cause acute,  
 CC localised infections such as otitis media, sinusitis and bronchopulmonary  
 CC infection and life-threatening, systemic diseases including endocarditis



Qy 661 GlyPheasPleuserThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsn 680  
 Db 82377 GGTTCGATGATTATCCACACCGACCTAGAGCATGATTAACCGAGATTATATCCAAAT 82436  
 Qy 681 GlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAla 700  
 Db 82437 GGTGGCTGGAATCCCTACTTATTCATTTGGCTTATGCTTGTGAATGAAGCTTAAACCCAAAC 82496  
 Qy 701 ProProGluTrpTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlu 720  
 Db 82497 CCCCCCAACATGCGAGGATTTTACCACTGATTTTGTCAATGGTAAGCCAAAGCCAAAGAG 82556  
 Qy 721 AlaLeuAlaIleAlaArgAlaIleValHisLysThrValAlaAspAsnLeuValAsnProMet 740  
 Db 82557 GCGTTATGGCAGGTGGTGGCTGCATTAACCGGTTCAGATTAATTTGTTGAATCCGATG 82616  
 Qy 741 ArgGlyIlyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAla 760  
 Db 82617 CGTGGCATCGTCAGCGCATTTCTTTAGAGGTGGCTCAACCGGTTCGATCGGATGCT 82676  
 Qy 761 AsnMetAlaIleAlaArgAlaIleValHisLysThrValAlaAspAsnLeuValAsnProMet 780  
 Db 82677 AATATGGCTATGCTCCAGCGGTATTAAGTGGCTGTATAGTTTGGCGATTAATGCTTAT 82736  
 Qy 781 GlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
 Db 82737 GCGCAGCATCGTCCCATACATGATGCTGGCATACAAAGCAGATTCATTTGGTCGGAT 82796  
 Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820  
 Db 82797 AATTATATCATGTCGCATATCGTTTGGCTTTTGGCTGGCGCAAGATTTGCT 82856  
 Qy 821 GlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGln 840  
 Db 82857 GGAATATGCACATGACAGTTTATCACTATATCAGATTAAGGTTATCTACAGCGCGTAA 82916  
 Qy 841 ValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860  
 Db 82917 GTATTGGCGGTGGTACAGCTGAATATATGAAATTAAGAAAGATTTCCGTTTGGCG 82976  
 Qy 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880  
 Db 82977 GTTTTGGTATATGTAATGCTTATGATTAAGGCTTTACTAATGATTAACAAATTTGCT 83036  
 Qy 881 AlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArgValAspValAlaThr 900  
 Db 83037 GCAGCTGTCGCTGTCCTGGCGCATCACTGTCGTCGAAGTTCGTTGATGTGCACT 83096  
 Qy 901 GlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 83097 GGTCTCAAGAGAGGCAATCCATTAAGCTGCTTTTATTTATGGCACACCATTT 83153

RESULT 3  
 AAAS0537  
 ID AAAS0537 standard; DNA: 2670 BP.  
 AC AAAS0537;  
 XX  
 DT 05-DEC-2000 (first entry)  
 DE Moraxella catarrhalis BASB081 mature protein coding region.  
 XX  
 KW BASB081 gene; infection; otitis media; pneumonia; sinusitis;  
 KW inflammation; therapy; antibacterial; antiinflammatory; vaccine;  
 KW diagnosis; ds.  
 XX  
 OS Moraxella catarrhalis.  
 XX  
 FH key location/Qualifiers  
 FT conflict replace(1003,G)  
 FT /tag- a  
 XX /note- "corresponds to base 1093 of AAAS0536"

PN W0200052042-A1.  
 XX  
 PD 08-SEP-2000.  
 XX  
 PF 23-FEB-2000; 2000WO-EP01468.  
 XX  
 PR 26-FEB-1999; 99GB-0004559.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 P1 Ruelle J;  
 XX  
 DR WPI: 2000-587296/55.  
 DR P-PSDB: AAY95988.  
 XX  
 PT New BASB081 polypeptides from Moraxella catarrhalis and polynucleotides  
 PT encoding the polypeptides used for treating infections, or as a vaccine  
 PT for preventing infections, especially those caused by M. catarrhalis -  
 XX  
 PS Claim 8; Page 60-61; 97pp; English.  
 XX  
 CC The present sequence is that of DNA encoding BASB081 mature protein  
 CC (see AAY95988) of Moraxella catarrhalis Mc2931 (ATCC 43617). It was  
 CC obtained by sequencing PCR-amplified DNA from a BASB081 expression  
 CC plasmid. The sequence shows 99.3% identity to a mature BASB081  
 CC coding sequence (see AAAS0537) obtained from M. catarrhalis ATCC 43617.  
 CC The invention provides BASB081 polypeptides, polynucleotides,  
 CC expression vectors, host cells, and a process for producing a BASB081  
 CC polypeptide. Also provided are vaccine compositions comprising a  
 CC BASB081 polypeptide or polynucleotide, and optionally at least 1  
 CC other M. catarrhalis antigen. A method for diagnosing a M.  
 CC catarrhalis infection involves identifying a BASB081 polypeptide, or  
 CC an antibody that is immunospecific for it, in a sample. A  
 CC therapeutic composition useful in treating M. catarrhalis diseases  
 CC in humans comprises an antibody directed against a BASB081 protein.  
 CC The disease can be a bacterial infection, e.g. otitis media in  
 CC infants and children, pneumonia in elderly, sinusitis, nosocomial  
 CC infections and invasive diseases, chronic otitis media with hearing  
 CC loss, fluid accumulation in the middle ear, auditive nerve damage,  
 CC delayed speech learning, upper respiratory tract infection, and  
 CC inflammation of the middle ear.  
 XX  
 S0 Sequence 2670 BP; 757 A; 527 C; 632 G; 754 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0 Length: 2670  
 Score: 4581.00 Matches: 888  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.89% Mismatches: 0  
 Query Match: 96.91% Indels: 0  
 DB: 21 Gaps: 0

US-09-914-168-2 (1-919) x AAAS0537 (1-2670)  
 Qy 31 GlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAlaIleAsn 50  
 Db 1 CAACAAATTAACCCCTGCAGCAATCATATGATACCCGCTCAAGACACCGCCATCAT 60  
 Qy 51 GlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeu 70  
 Db 61 CAAGCAAGAGCAGCGCAATCCGCTGTTTGCACACCTGACAGATTAACACACGCTT 120  
 Qy 71 AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnAsp 90  
 Db 121 AATGCTGCTGGACGTAAGCTTAAGCCCAATCAACAGCTTGGATGTGTCATTTTAT 180  
 Qy 91 AspGlnSerProIleSerArgIleGlyLysGlnSerProProLeuGlyLeuAspMetSer 110  
 Db 181 GATCAATCGCCGATATCTCGATATCGTGAGCAATCAACCCCTTGGTTGATATGTCG 240  
 Qy 111 ValIleGluGluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGlu 130  
 Db 241 GTCATCGAAGAAACACACCGCTTAAGCTTGGAGGAATTAATTTGTCAGAAATCTACTGAG 300

OY	131	MetGlyIleasnProasnAspTyrIleProgluTyrGlnGlyIuclnProasnSerGlu	150
Db	301	ATGGGAATCAATCCAAATGATTAATTATTCAGAAATATCAAGCGCAACCTAATAGTGAG	360
OY	151	ValValValProProthbPheuGluProGluLysProGlyLeuIleLysArgLeuTyrAla	170
Db	361	GTGGTTTACACACCGACATTAGAACTTGAAAAACAGGTTTGATCAAGGCTCTTATGCA	420
OY	171	ArgLeuPheasnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSer	190
Db	421	CGCCATTATTATGATGTGTGTCATAAAGTGCCTTACGCTTAAGSCAAATTTTATCATCA	480
OY	191	SerGlnSerArgLysIleThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla	210
Db	481	TGCGATATAGCGGAACCAAGTGGCATTTGGTGTATCCATCAAAAAACAGACGCTTAATGA	540
OY	211	AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuasnGlySer	230
Db	541	AATATCAAAACGACACTTGAAGACATCACCCAGACGTCAAGCATTTGATTAAGTCTCT	600
OY	231	IleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle	250
Db	601	ATCCACCCCTTAAGGCAACGTCTTGGTGGACGCGTGTCTGCTTATTTATCATATT	660
OY	251	AspLeuSerIleIleLeuAsnSerIleGlyGluValAspValIleIleHisAspLeuGly	270
Db	661	GATTATCATCATAAACAATAATACATCGAGAGAGCTGGATCTCATCTCATCATATTAGCT	720
OY	271	GluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys	290
Db	721	GAACCTGTTTATTTGATTTATCGACGCGTGAGAGCTACGAGGTGAAGGTCTCATATATAA	780
OY	291	AlaPheThrTyrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGly	310
Db	781	GCATTTACTACCGTGGCGGATGAGGTGCCATTCCTCATGGCGATGCTTTCATACAGCG	840
OY	311	LysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAsp	330
Db	841	AAGTACGAACCAAAAAAATCTCATCGAANAATGCCAGTGTGAACATGATATTGTAT	900
OY	331	GlyLysArgPheLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSer	350
Db	901	GGCGGTTGCGGATTCCTTCAGTTGATGTAATTTTCCAGATTAATAACCGCTCATGTCAGC	960
OY	351	LeuIleTyrAspThrArgLysThrGlnTyrArgPheAspGluValValPhePheThrIleAsp	370
Db	961	TTAATTTATGATACAGGTACGCAAGTATCGCTTTGATGACAGCTGATTTTTTACCTTAT	1020
OY	371	ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeu	390
Db	1021	CCTAAACCAATCAATTGACAAACCGATCCAGATTAAGCTCCAGTTAAACGAAATTACTT	1080
OY	391	GluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGluAlaValAlaArgAlaLeu	410
Db	1081	GAGCAGTTACTCACCCCTTAACATGATGGAGAGAGCTTACAATTATACAGCGGTGCGTCACTT	1140
OY	411	SerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhePro	430
Db	1141	TCAATATGATTGATTGGCCACACGATTTTAAATATGTAATAACGACATTTGCTTTTCCA	1200
OY	431	GluArgGluGlnIleGlnAsnAspGluValSerPheGluGlnLysSerSerSerArgThr	450
Db	1201	GAGCGTACACGATCCAAAAACGACCAAGTAGCTTTGACACACTTTTCAAGTACGCGTACT	1260
OY	451	GluProArgGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThr	470
Db	1261	GAACACGACACAGTGTGATGAACACACTTGAACCTGTATTGAACACGCTTAGACTAAGC	1320
OY	471	AspGlyIleLeuMetAspLysSerProIleGluPheSerAlaSerAsnLeuIleGlnAsp	490
Db	1321	GATCGGATTAATTAATGATATTTCGCCATTCGAATTTAGCATCTAAATCTATTTACAGAC	1380

QY	491	LysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTrpAspMetProAspAspArgVal	510
Db	1381	AAAGCTAAATTTGGTGGCCCAAGAGCGCCGACATTATATACATGGCTGTGATAGGGTG	1440
QY	511	LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAla	530
Db	1441	CTTGGCATCATCATGATGATGAGGGGTAAATCGCTCTATTTTGGGAGGAAATCACCGATGCC	1500
QY	531	ValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeu	550
Db	1501	GATATGCCGTTGGTGCACGGCTATTATTACCGATGAATCATGAAATAGAGTAATAGATTGG	1560
QY	551	ProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValIyrGlnSerLysLys	570
Db	1561	CCCGAGCGTACCGCATTTGGCTAAATCGACAGACCCCTGCTGATGCTATCAAGATGAAAAA	1620
QY	571	ValProLeuTrpValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGly	590
Db	1621	GTGGCGCTATATGTCCTTTGGTGGAGTGATAAACCCGAGATGTCAAATTTGGTTTGGCG	1680
QY	591	TrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg	610
Db	1681	TGGGATGAGACACAGGATACCGGCTATGTCACAAAATTTGACATATTTGATTATATCGT	1740
QY	611	AspGlyTrpGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValAlaLysLeu	630
Db	1741	GATGGCTATCAAGCAGGGGGGAGACTAAAGCTGTGTGAGATATAAAAAAGGGGTCAAAATT	1800
QY	631	TyrAlaThrLysProLeuSerHisProLeuAsnAspGluLeuArgAlaThrLeuGlyTrp	650
Db	1801	TATGCCACCAAAACCGCTTAGCCACCCCTCTAAATGATACGTAAAGACCAACTTTGGGTAT	1860
QY	651	GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu	670
Db	1861	CAACAGAGAGTTTTGGTGCACCTGACCAATAGTTTTGATTATCCACAGCACCTAGAG	1920
QY	671	HisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTrpSerLeuArg	690
Db	1921	CATGGATTACCCGACAGATTATATCCAAAAGTGGCTGATCTGCTATTATTCATTGGCT	1980
QY	691	TyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGluAspLeuProVal	710
Db	1981	TATCGTCTGTATAGCTTAAACCCCAAGCACCCCTGAAACATGACGAGATTACCACTG	2040
QY	711	AspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys	730
Db	2041	GATTTTGTCAATGGTAAAGCCCAAGCAAGAGCGCTTTTGGCAGAGTGTTCTGTCAATAA	2100
QY	731	ThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTrpSerLeuGlu	750
Db	2101	ACGGTGCAGATATTGGTTGTTAATCCATGCGTGGCTATGCTGACGAGATATCTTAAGG	2160
QY	751	ValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSer	770
Db	2161	GTTGGCTCAACCGCTTGTGTATCGGATGCTAAATATGGCTATTGCTGCAGACTGTTATAGT	2220
QY	771	GlyValIyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly	790
Db	2221	GGCGGTATAGTTTTGGGGAATAATGCTTAAGCAGCAATCGTCCCATGATGACTGGT	2280
QY	791	GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTrpArgLeuArg	810
Db	2281	GGCATACAGAGAGATATCTTTGGTGGGATTAATTTAATCATCATGTCGCAATCGTTTGGCT	2340
QY	811	PhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIle	830
Db	2341	TTTTTCTGCTGGCGACCAAGATTTCGTGGATATGACACATGATTAATTCACACTATA	2400
QY	831	SerAspLysGlyTyrLeuThrGlyGlnGlnValLeuAlaValGlyThrAlaGluIyrTrpAsn	850
Db	2401	TCAATATAGGCTTATCTGCACAGCGCGCTCAAGTATTTGGCGCTTGATACGCTGAATATAT	2460
QY	851	TyrGlnPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTrpAsp	870

|||||  
 Db 2461 TATGAATTTATGAAAGATTTCGCTTGGCGGTTTGGTGAATATGTGAATCTTATGAT 2520  
 Qy 871 LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValArgTyrPalaSerPro 890  
 Db 2521 AAAGGCTTACTAATGATACCAAAATTTGTGTCAGGTGTGGTGGTGGCATTACACT 2580  
 Qy 891 ValGlyGlnValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLys 910  
 Db 2581 GTGGGTCAAGTTCGTGTGATGTGGCAACTGTGTCAAGAAGAGGCAATCCATTAAAG 2640  
 Qy 911 LeuHisPhePheIleGlyThrProPhe 919  
 Db 2641 CTCGATTTTATTTATGGCACACACATTT 2667  
 RESULT 4  
 AAF61044  
 ID AAF61044 standard; DNA: 1374 BP.  
 AC  
 XX AAF61044:  
 XX 16-MAY-2001 (first entry)  
 DE P. putida KT2440-associated DNA ORF06604.  
 KW Transgenic plant; detection; probe; amplification; vaccine carrier;  
 KM microbial production strain; biological remediation; ds.  
 XX Pseudomonas putida.  
 OS  
 XX DE19935088-A1.  
 PN 01-FEB-2001.  
 PD 27-JUL-1999; 99DE-1035088.  
 PF 27-JUL-1999; 99DE-1035088.  
 PR 27-JUL-1999; 99DE-1035088.  
 XX (TIGR-) TIGR INST GENOMIC RES.  
 PA (OUIA-) OUIAGEN GMBH.  
 PA (GBFB-) GES BIOECHANOLOGISCHE FORSCHUNG MBH.  
 PA (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.  
 PA (MED1-) MEDIZINISCHE HOCHSCHULE HANNOVER.  
 XX  
 DR WPI: 2001-192469/20.  
 XX  
 PT New DNA sequences specific for Pseudomonas putida KT2440, useful as  
 PT genetic engineering host, allow detection in presence of other  
 PT related bacteria .  
 XX  
 XX Claim 1a: Page 94; 158pp; German.  
 PS  
 XX  
 CC This invention describes novel DNA sequences (I) for specific detection  
 CC of Pseudomonas putida KT2440. The invention also describes (1)  
 CC recombinant expression vector containing (1); (2) prokaryotic or  
 CC eukaryotic cells transformed or transfected with (1) or the vector of  
 CC (1); (3) production of expression products by culturing cells of (2);  
 CC (4) expression products, or their fragments, of (1) and synthetic  
 CC proteins or peptides with the same sequences (A); (5) poly- or  
 CC monoclonal antibodies (Ab) that react specifically with (A); (6)  
 CC hybridoma cells that produce the monoclonal Ab of (5); (7) transgenic  
 CC plants that contain transformed or transfected cells of (2); (8)  
 CC detecting KT2440 using a labeled (1) or Ab as probe; and (9) DNA chips  
 CC carrying one or more (1). (1), and their fragments, are used as probes  
 CC to detect and isolate full-length cDNAs and/or to amplify such cDNAs by  
 CC polymerase chain reaction, and for production of transgenic plants. (1),  
 CC or antibodies that recognize their expression products, are used for  
 CC detecting the presence of KT2440, particularly in presence of other,  
 CC even closely related, bacteria. KT2440 is one of the bacteria classified  
 CC work, e.g. as microbial production strains, for genetic engineering  
 CC and as vaccine carriers. (1) are exclusive to KT2440 with no significant  
 CC homology with sequences in other bacteria (specifically the closely

CC related pathogen P. aeruginosa). Compared with other 'safe' bacteria, it  
 CC has greater catabolic activity and better survival in, and adaptation to,  
 CC the rhizosphere and soil.  
 XX  
 SQ Sequence 1374 BP; 264 A; 428 C; 429 G; 251 T; 2 other; \*  
 Alignment Scores:  
 Pred. No.: 6,84e-40 Length: 1374  
 Score: 604.50 Matches: 165  
 Percent Similarity: 40.41% Conservative: 90  
 Best Local Similarity: 26.15% Mismatches: 195  
 Query Match: 12.79% Indels: 181  
 DB: 22 Gaps: 12  
 US-09-914-168-2 (1-919) x AAF61044 (1-1374)  
 Qy 290 LysAlaPheThrValAlaAspGluValProLeuIleGlyAspValPheHis 309  
 Db 4 AAGGCTTC---CGCGTCCGTGACAGCAGCGCGCTGCGCGGTGACACACTCAATCAC 60  
 Qy 310 GlyLysTyrGluThrLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329  
 Db 61 GGCATTACKAGGATGCAAGCGGTGATCAAGAACCGAGCGTCGCGTATGGCTTCTTC 120  
 Qy 330 AspGlyArgTyrPLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349  
 Db 121 AGTGGCGGTTCAGTACAGCAGCGCTGCGCCGTGACGCCCAAGCGCGTGTGCGATATC 180  
 Qy 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIle 369  
 Db 181 GAACGTGCTTACAGAGAGGCGCGGTATCCGTCGGCGCGCTCACCTTC----- 231  
 Qy 370 AspProLysThrAsnGluLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389  
 Db 232 -----GGTGGCGACACACCGCTGACGAGACCTG 261  
 Qy 390 LeuGluGlnLeuLeuThrValAlaSerMetGlyGluAlaTyrAsnLeuGlnAlaValArgAla 409  
 Db 262 CTCGACGCGATGTCGTGTCGTTCAAGCGCGGTACCCCTTACAGCTCGGAACTGTCGACAG 321  
 Qy 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429  
 Db 322 CTCGACACAGCAGCTCGAATCGAGCGGCTATTTTC----- 354  
 Qy 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449  
 Db 354 ----- 354  
 Qy 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469  
 Db 354 ----- 354  
 Qy 470 ThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln 489  
 Db 355 ---GAAGGCGTCGCGGTGATCGCGCGCC----- 381  
 Qy 490 AspLysLeuAsnLeuValAlaIleLysAlaArgHisLeuTyrAspMetProAspArg 509  
 Db 381 ----- 381  
 Qy 510 ValLeuAlaIleAsnHisAspArgGlyValAsnArgSerIleLeuGluArgIleSerAsp 529  
 Db 381 ----- 381  
 Qy 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAsp 549  
 Db 381 ----- 381  
 Qy 550 LeuProGluArgTyrThrAlaLeuAlaSerAsnArgLysThrProAlaAspValTyrGlnSerLys 569  
 Db 382 -----ACTGCTGCTGTGGCGCAGAGAGTCGCGGTGATGTCAT----- 420  
 Qy 570 LysValProLeuTyrValAlaPheValAlaSerAspLysProArgAspGlyGlnIleGluLeu 589

```

Db 421 -----CTGGAACCGGTAACCCAGCACCACCATGGCGCTTGCGCTG 459
OY 590 G1YTRPG1YSerAspThrG1YThrArgLeuVal1ThrIysPheGlnHisAsnLeuIleAsn 609
Db 460 GCGTTCTCGACCGAGCTGCGCGCGCGCGCAAGCCAACTGGACCCCGGCTGCGTCAAC 519
OY 610 ArgAspG1YTrgGlnAlaG1YAlaG1YLeuArgLeuSerGlnAspIysGlyValIys 629
Db 520 CCACAAAGCCACAGCTATGCTGGCAACCCAACTGCTGCGCGCGCGCAAGCTGCGC 579
OY 630 LeuYrAlaThrIysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGly 649
Db 580 CTGGGTATGACATTCTCCCTCGACCCCGCTGACCCGCAAGTTCCTTGGCGCGCGC 639
OY 650 TyrGln--GlnGluVal1PheG1YHisSerThrAsnG1YPheAspLeuSerThrArgThr 668
Db 640 TACCGAAGCAGAGAGCTTCCGCGCGACACACCGCTACACAGCTTTCACGTCGCGCCC 699
OY 669 LeuGlnHisGluIleSerArgSerIleIleGlnAsnG1YTrpAsnArgThrTyrSer 688
Db 700 GAGTCGACACACCAAGCTGCCCTG-----GGCTGGCAGCGCGCTGATTTCG 744
OY 689 LeuArgTrgArgLeuAspIysLeuIysThrGlnAlaProProGlnIuThrTrpGlnAspLeu 708
Db 745 CTCAGATCCAAAGCGGAGAAATATGCTG-----774
OY 709 ProValAspPheValAsnG1YIysProSer-----GlnGluAlaLeuLeuAlaGlyVal 726
Db 775 -----GGTGACGACTCCGCTTGGACCAACCTGCTGATCGCGGCGCTC 816
OY 727 AlaValHisIysThrValAlaAspAsnLeuValAsnProMetArgGlyTrgArgGlnArg 746
Db 817 AGCTTTTCTCTTCCTCGCCAGGATGACACAGTATCGATCGGCAACAGCGCTATCGCTGAC 876
OY 747 TyrSerLeuGluValAlaG1YSerSerG1YLeuVal1SerAspAlaAsnMetAlaIleAlaArg 766
Db 877 TTCATGTTCCAAAGTCCCAAGAGAGCGTGTGTCGACACCAACCTGCTGACGCGCAC 936
OY 767 AlaGlyIleSerGlyVal1YrSerPheG1YAspAsnAlaTrgIySerAsnArgAlaHis 786
Db 937 GTATTGCTCAAGGCGCTGACCACTCGGCACAAAC-----CAC 975
OY 787 GlnMetThrGlyGlyIleGlnAlaGlyIyIleTrpSerAspAsnPhe--AsnHisVal 805
Db 976 CGCTTTCTCGGACCGGATCTTTCGCGGTGGCAGTGCACCAACCAATGGCTTCACAGACAACTT 1035
OY 806 ProTrgArgLeuArgPhePheAlaG1YIysAspGlnSerIleArgG1YTrgAlaHisAsp 825
Db 1036 CCGCGCTCGCTGCGTTTCTTTCGCGGTGGCAGTGCACCAAGTGTGCGGTACGACTACAG 1095
OY 826 SerLeuSerProIleSerAspIysG1YTrgLeuThrGlyIysGlnVal1LeuAlaValGly 845
Db 1096 ACCCTGTCGCGAAGAACAGCAGCAGCTGATCGCGCGGCTACTTGTGTCGACAGC 1155
OY 846 ThrAlaGluTrgAsnTrgGlnPheMetIysAspLeuArgLeuAlaVal1PheG1YAspIle 865
Db 1156 AGTGTGCGATACAGATATCTCTGACGAAATAATGGCGGCTGACGCTGCTGCAACAA 1215
OY 866 G1YAsnAlaTrgAspIysG1YPheThrAsnAspThrIysIleGlyAlaG1YValGlyVal 885
Db 1216 GGCACACTCGTTCAAGAGCTGAGCTGCGCCAGCGCTCAAGACGCGGTGCTGGTGGTGTG 1275
OY 886 ArgTrgAlaSerProValG1YGlnValArgValAspValAlaThrGlyValIysGlnGlu 905
Db 1276 CGCTGGGTATGCGCAAGTGGCGCGCTGCGCTGACGCTGCGCAAGGCGCTGATGACGAA 1335
OY 906 G1YAsnProIleIysLeuHisPhePheIleGly 916
Db 1336 GCGGCGC---ATTGCGCTGCACTTTTTCATGGCG 1365

```

```

ID AAA50269 standard; DNA; 1737 BP.
XX
AC AAA50269;
XX
DT 07-NOV-2000 (first entry)
XX
DE Haemophilus influenza strain Rd KW20 BASB067 gene.
XX
KW BASB067 gene; outer membrane protein; vaccine; antibiotic;
XX antibacterial; screening; infection; diagnosis; therapy; ds.
XX
OS Haemophilus influenzae.
XX
FH Key Location/Qualifiers
FT CDS 1..1737
FT sig_peptide 1..66
FT mat_peptide 67..1734
FT /*tag= c
XX
PN W0200047737-A1.
XX
PD 17-AUG-2000.
XX
PE 04-FEB-2000; 2000WC-EP00887.
XX
PR 09-FEB-1999; 99GB-0002880.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J, Thonnard J;
XX
DR WP1: 2000-515059/46.
DR P-PSDB: AAY95820.
XX
*PT BASB067 polypeptide and polynucleotide from Haemophilus influenzae are
PT used for diagnosing and treating H. influenzae infections -
XX
PS Claim 3; Page 77-78; 87pp; English.
XX
CC The present sequence is that of the coding region of the BASB067
CC gene of Haemophilus influenzae strain Rd KW20. It encodes a
CC 578-amino acid surface expressed protein (see AAY95820) that is
CC recognized by the immune system. The invention relates to
CC recombinant materials and methods for the production of BASB067
CC polypeptides and polynucleotides, for use especially in therapeutic
CC and prophylactic vaccines. It also relates to methods for using
CC such polypeptides and polynucleotides in the prevention and
CC treatment of microbial diseases, in diagnostic assays for detecting
CC diseases associated with microbial infections, and assays for
CC detecting expression or activity of BASB067 polypeptides or
CC polynucleotides. A polynucleotide having at least 85% identity
CC to the present sequence can be used in the recombinant production
CC of BASB067 immunogenic polypeptides in transformed host cells, and
CC in vaccine compositions.
XX
SQ Sequence 1737 BP; 517 A; 309 C; 380 G; 531 T; 0 other;
XX
Alignment Scores:
Pred. No.: 4.6e-26 Length: 1737
Score: 436.50 Matches: 142
Percent Similarity: 35.64% Conservative: 105
Best Local Similarity: 20.49% Mismatches: 235
Query Match: 9.23% Indels: 211
DB: Gaps: 14
US-09-914-168-2 (1-919) x AAA50269 (1-1737)
OY 243 ArgAlaValGlyTrgYrAspIleAspLeuSerIleIleArgAsnSer---IleGlyGlu 261
Db 217 CGTGTGTTGGTTATATGAAATCTTCGCTGCTTTCAGACGAAACAGCGTCAAGCGCAA 276

```

```
OY 262 ValAspValIleIleHisAspLeu-----GlyGluProValIleAspPheArgAla 279
    |||:|||||
Db 277 CGCATTTATTGATTGATTCATGACACCAACGCGGCGGAGCTGAT 336
OY 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluVal 299
    |||:|||||
Db 337 GTCGCAATTTAGGGGAGCGCGCACAGATGAAATTTTAAATGATCCGTAACAACTTA 396
OY 300 ProLeuLeuIleGlyAspValPheHisIleGlyLysTyrgluThrLysAsnLeuIle 319
    |||:|||||
Db 397 CCG---AAAGATGCCGTTTGGTGGAGCACCAACTTACGATGATACAAACAGCGCAT 453
OY 320 GluAsnAlaSerAlaGluHisGlyTyrrPheAspGlyArgThrPheLysAspArgSerValAsp 339
    |||:|||||
Db 454 TCACGTTTACGATTAATTCGTGGGATTTTGTATGGAGACTTAAATTTTACCGTTTAA 513
OY 340 ValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrrAspThrGlyThrGlyIntyr 359
    |||:|||||
Db 514 ATCAGCCCTGAAACCCATCAAGCATGTGGCGAATGTATTTGATGATGTCCTCAT 573
OY 360 ArgPheAspGluValValPhePheThrIleAspProLysThrAsnGluLeuThrThrAsp 379
    |||:|||||
Db 574 CATTATGGCAATATTACTTT----- 594
OY 380 ProAspLysLeuProValIleArgGluLeuLeuGluLeuLeuThrValAsnMetGly 399
    |||:|||||
Db 595 ---AGCCATTACAAATTTCCGATGATCTTAATATATATTCTTAACATCAATCTGCT 651
OY 400 GluAlaIleTyrrAsnLeuGluAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr 419
    |||:|||||
Db 652 GATCCGATTTAAATGAATTAATTTGTCGATTTAACACAGCGATTTCCATCTTCAAAATGG 711
OY 420 PheAsnMetValAsnThrGluIleValPheProGluArgGluGluIleGluAsnAspGln 439
    |||:|||||
Db 712 TTATAGTCACAGA----- 723
OY 440 ValSerPheGluGlnSerSerSerArgThrGluProAlaGluValAspGluSerThr 459
    |||:|||||
Db 723 ----- 723
OY 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479
    |||:|||||
Db 723 ----- 723
OY 480 IleGluPheSerAlaSerAsnLeuIleGluAspLysLeuAsnLeuValAlaAlaLysAla 499
    |||:|||||
Db 724 -----TAGTTCAGCCCTAATGTTAAT----- 744
OY 500 ArgHisLeuTyrrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
    |||:|||||
Db 744 ----- 744
OY 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539
    |||:|||||
Db 744 ----- 744
OY 540 ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
    |||:|||||
Db 744 ----- 744
OY 560 LysThrProAlaAspValIleTyrrGlnSerLysLysValProLeuTyrrValPheValAlaSer 579
    |||:|||||
Db 745 -----CATAAAGCAAACTGTGATGTGAGATTAATTTATTCATCA 786
OY 580 AspLysProArgAspGlyGlnIleGlyLeuGlyIleArgLysSerAspThrGlyThrArgLeu 599
    |||:|||||
Db 787 CGTAAATAAAATGCGATGGAACCTGCTGGCTTTCTACTGATGGCGGCTTCACGGA 846
OY 600 ValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrrGlnAlaGlyValAlaGluLeu 619
    |||:|||||
Db 847 CAATATAGCTGACAAACCTTGATTAATAGCCGTGACACATAGTTCCGTTCAAAATCTT 906
OY 620 ArgLeuSerGluAspLysLysGlyValLys--LeuTyrrAlaThrLysProLeuSerHis 638
```

```
Db 907 TATCTCTCGACCAAAACAACTCTAGAGGCACTATGCAATGCACCTGTTAAATAT 966
OY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrrGlnGlnGluValPheGlyHisSer 658
    |||:|||||
Db 967 CCATTAAATTTATTACTATGATTTTCCCGTGGTGGGA-----GGGGAATAA 1014
OY 659 ThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIle 678
    |||:|||||
Db 1015 GAGAAAT-----GATACCAATTCAGAGAGTGTCTACGTTGTCAGCGTTATTTGGAAT 1068
OY 679 GluAsnGlyGlyTyrrAsnArgThrTyrrSerLeuArgTyrrArgLeuAspLysLeuTyrrThr 698
    |||:|||||
Db 1069 AATGCGCATGGTTGSCAATATTTTGGCGAGCTTCGTATGCGATACGACAGTTT---ACA 1125
OY 699 GlnAlaProProGluThrThrGlnAspLeuProValAspPheValAsnGlyLysProSer 718
    |||:|||||
Db 1126 CAGCGGAT----- 1134
OY 719 GlnGluAlaLeuLeuAlaGlyValAlaHisLysThrValAlaAspAsnLeuValAsn 738
    |||:|||||
Db 1135 -----ATCAGTCAATAAACCTTA-----CTTCTTTAT 1161
OY 739 ProMetArgGlyTyrrArg----- 744
    |||:|||||
Db 1162 CCAACTGTTGATTTACTGCTACTGATTAACGTGATGCGTTCCTTGGCAGTTGGGCGCAT 1221
OY 745 ---GlnArgTyrrSerLeuGluValGlyLysSerGlyLeuValSerAspAlaAsnMetAla 763
    |||:|||||
Db 1222 GTGCAAAATAATTAATCTTTGATTAAGCAAAACGAATTTGGCTVTCAGAAATCTTTTATA 1281
OY 764 IleAlaArgAlaGlyIleSerGlyValTyrrSerPheGlyAspAsnAlaIleGlySerAsn 783
    |||:|||||
Db 1282 AAAGTCAAGCATCTAAGCGCGGTGCTGCTACTTATGCGAATAAT----- 1326
OY 784 ArgAlaHisGlnMetThrGlyGlyIleGluAlaGlyTyrrIleTrpSerAspAsnPheAsn 803
    |||:|||||
Db 1327 -----CATCGTGTGCTGCTGCTGCTGAAATCGGGTATTTACATACAAAGCTATTGCA 1380
OY 804 HisValProTyrrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrrAla 823
    |||:|||||
Db 1381 AAAATTCGCGCTACACTGCTGCTTTTCTGCTGGCGATGCTGAGTGTGCGGTTACGGC 1440
OY 824 HisAspSerLeuSerProIleSerAspLysGlyTyrrLeuThrGlyGlyGlnValLeuAla 843
    |||:|||||
Db 1441 TATATAAAATTTGCCCTTAATAATAGAAATGCAAAATTTGCTGCGCATTTGCTT 1500
OY 844 ValGlyThrAlaGluTyrrAsnTyrrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
    |||:|||||
Db 1501 ACCACTTCTTTAGATATATCAATATATCAAGTCTATCCAAATTTGTTGGCGGCACACTTTTGA 1560
OY 864 AspIleGlyAsnAlaIleTyrrAsnTyrrGlyPheThrAsnAspThrLysIleGlyValAla 883
    |||:|||||
Db 1561 GATAGTGAATTAAGCTGCGCATTAATTAACACAGCAAAAGACTCTGTTATGCGACAGCGCTT 1620
OY 884 GlyValArgTyrrAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys 903
    |||:|||||
Db 1621 GGTGTGCGTTGGCATTCGCCAATTTGTGCGATTAATTTGATATTTGCCACACCATTCCT 1680
OY 904 GluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
    |||:|||||
Db 1681 GATTAAGATTAACAGCAAAATATTTCAATTTTACATCCGA 1719
RESULT 6
ID AAA50270 standard; DNA; 1731 BP.
AC AAA50270;
XX
XX
XX 07-NOV-2000 (first entry)
XX
XX Haemophilus influenza non-typable strain 289 BASB067 gene.
DE
```

KW BAS067 gene; outer membrane protein; vaccine; antibiotic;  
 KW antibacterial; screening; infection; diagnosis; therapy; ds-  
 XX  
 XX Haemophilus influenzae.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 1..1731  
 FT /\*tag= a  
 FT sig\_peptide 1..60  
 FT /\*tag= b  
 FT mat\_peptide 61..1728  
 FT /\*tag= c  
 PN MO200047737-A1.  
 PD 17-AUG-2000.  
 PF 04-FEB-2000; 2000MO-EP00887.  
 PR 09-FEB-1999; 99GB-0002880.  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PI Ruelle J, Thonnard J;  
 DR WPI: 2000-515059/46.  
 DR P-PSDB: AAY95821.  
 PT BAS067 polypeptide and polynucleotide from Haemophilus influenzae are  
 PT used for diagnosing and treating H. influenzae infections -  
 PS Claim 3; Page 80; 87pp; English.  
 XX  
 CC The present sequence is that of the coding region of the BAS067  
 CC gene of Haemophilus influenzae non-typable (Nth) strain 289. The  
 CC DNA was isolated by PCR using primers (see AAA50271-72) specific for  
 CC the BAS067 gene. It encodes a 576-amino acid surface exposed  
 CC protein (see AAY95821) that is recognised by the immune system. The  
 CC invention relates to recombinant materials and methods for the  
 CC production of BAS067 polypeptides and polynucleotides, for use  
 CC especially in therapeutic and prophylactic vaccines. It also  
 CC relates to methods for using such polypeptides and polynucleotides  
 CC in the prevention and treatment of microbial diseases, in diagnostic  
 CC assays for detecting diseases associated with microbial infections,  
 CC and assays for detecting expression or activity of BAS067  
 CC polypeptides or polynucleotides. A polynucleotide having at least  
 CC 85% identity to the mature portion of the present sequence  
 CC (nucleotides 61-1731) can be used in the recombinant production  
 CC of BAS067 immunogenic polypeptides in transformed host cells,  
 CC and in vaccine compositions.  
 XX  
 SQ Sequence 1731 BP; 509 A; 322 C; 386 G; 514 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 2,48e-25 Length: 1731  
 Score: 427.50 Matches: 141  
 Percent Similarity: 35.64% Conservative: 106  
 Best Local Similarity: 20.35% Mismatches: 235  
 Query Match: 9.04% Indels: 211  
 DB: 21 Gaps: 14  
 US-09-914-168-2 (1-919) x AAA50270 (1-1731)  
 QY 243 ARGALAVAGLYTYRFRASPILEASPLEUSERILEIARGASNER---ILEGLYGIU 261  
 DB 211 CGTGTGCTTATTTATGATCTCCGTCGTTTGAACGAAACAGCGTCAAGGCATA 270  
 QY 263 VALASPVALLIEILEHISASPLEU-----GLYGIUProvalTYRILEASPTYRARGALA 279  
 DB 271 CGCGATTATTCATGTCATGTACACGAGCGAGCAACAAATAATGCGGAGCTGAT 330  
 QY 280 VALGIUVALARGGLYGLIYALASPSAPLYSALAPHERHTHVALALASPLUVAL 299  
 DB 961 CCATTAAATTATTAATGATTTTCCGCTGGGAA-----GGGGAATAA 1008

DB 331 GTGCAATTGAGGGGAGAACCCGACACAGATGAAAAATTTGATGCGCTACGTAACCTTG 390  
 QY 300 PROLEULENILEGLYASPVALLPHEHISISGLYTYRGLUHTHLYSASLEULE 319  
 DB 391 CCA---AAGAGAGCGCTTTGTTGACACCAACCTACATGATTAACAAACAGCGATT 447  
 QY 320 GLUASNALASERIALAGLHISISGLYTYRPHASRGLYATGTRPLEUASPRASERVAlASP 339  
 DB 448 TCACGCTTGCGATTAAATCGTGATTTTGTATGATGGAACCTTAAATTCACGTTAGAA 507  
 QY 340 VALILEUPROASPVASNTHTHALASPVASERLEULETYRASPHTHGLYTHGLTYR 359  
 DB 508 ATCAGCCCTGAACCCATCAACATGAGTGGCAGATGTTATTGATGAGTGCTCCGTTAT 567  
 QY 360 ARGPHASPRGLUVALVALPHERPHERHTHILEASPROLYSTHRASGLINLEUTHRASP 379  
 DB 568 CATTATGCAATATTAATCTT----- 588  
 QY 380 PROASPLYSLEUPROVALYSARGLULEULEUGLNULEULEUTHRVALASMETGLY 399  
 DB 589 ---AGCATTCACAAATCCGAGCATTTATCTCAATAATTTCTTAACATCAATCTGCG 645  
 QY 400 GLUALATYRASNLEUGLINALVALARGALALEUSERASNASPLEULEIATHTARGTYR 419  
 DB 646 GATCCATATTTTAATGAATAATTTGTCGATTTAAACAGCATTTTTCATCTCAAAATGG 705  
 QY 420 PHEASNMEVALASNTHTRGUILEVALPHERPROGLUARGUINILEGLINASNASPGLN 439  
 DB 706 TTTAGCTCAGTA----- 717  
 QY 440 VALSERPHEGLINSERSERSESRARGTHRGUPROALAGLNUVALASPLUSERTHR 459  
 DB 717 ----- 717  
 QY 460 LEUGLUPROVALILEGLUHTHVALGLULEUTHRASPOLYILEULEMETASPLIESERPRO 479  
 DB 717 ----- 717  
 QY 480 ILEGLUPHESERIALASERASLEULEINLASPLYSLEUNASLEUNVALALALATLSALA 499  
 DB 718 -----TTAGTTCAAGCCTAATGTTAAT----- 738  
 QY 500 ARGHISLEUTHYRSPMETPROASPRASPRAGVALLEUALALEASNHISASPSGLYVAL 519  
 DB 738 ----- 738  
 QY 520 ASNARGSERILELEUGLYARGLIESERASPVAlASERAlVALAlARGLAlAlELEU 539  
 DB 738 ----- 738  
 QY 540 PROASPLUSERGLUASGLNUVALILEASPLEUPROGLUARGTHRALALEUAlASNARG 559  
 DB 738 ----- 738  
 QY 560 LYSTHTRPROALASPVALLTYRGINSERLYSVALPROLEUTHYRVALPHEVALALASER 579  
 DB 739 -----CATAAAGCAAACTGTGAGTGTGAGATGATTTATCTTATCA 780  
 QY 580 ASPLYSPROARGASPOLYGLINILEGLYLEUGLYTRPGLYSERASPHHTHLYTHARGLEU 599  
 DB 781 CGTAAAAAANAATGCGATGACACTGCTGCGCTTTGCTACTGATGCGCGCTTCACGGA 840  
 QY 600 VALHTHLYSPHEGLUHHISASLEULEASNARGASPRGLYTYRGLNALAGLYALAGLUEN 619  
 DB 841 CAAATTTGGCTGACAAACCTTGATTTAATACCGTGACATAGTTTGGCTCAATATCT 900  
 QY 620 ARGLEUSERGLUASPLYSGLYVALYS---LEUTYRALATHLYSPLEUSERHTS 638  
 DB 901 TATCTCTCTGCAACCAAACTCTAAGAGCAACTTATGCAATGCCACTGCTTAAATAT 960  
 QY 639 PROLEASNASPGLINLEUARGALATHREUGLYTYRGINGLNUVALPHEGLYHISER 658  
 DB 961 CCATTAAATTATTAATGATTTTCCGCTGGGAA-----GGGGAATAA 1008





```

Db 1968 GCGAAATCGAACTGTCCTCCGCTCAG----- 1942
Oy 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356
Db 1941 CAGACCGCGTTTAACTGCTCATCATCATCATCGACGCG-----ACCGGT 1897
Oy 337 ThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeu 376
Db 1896 CCGAAACACATGACATGACAACTGACTGCGCAAGCCGTTGAGATGCTG 1837
Oy 377 ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrVal 396
Db 1836 GTAAACCGTTCATTGAG---CCGCTGAAAGTTGCATGACGACCGCTGCTCCGGA 1780
Oy 397 AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnLeuIleAla 416
Db 1779 TCTCAT-----ATCGACGACGTTATC--- 1759
Oy 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGln 436
Db 1758 -----CTCGTTGGTGGTCACTCGATGCCA-----ATGGTTGAG 1723
Oy 437 AsnAspGlnValSerPhe-----GluGlnSerSerSerSerArgThrGluProAla 453
Db 1722 AAGAAAGTTGCTGAGCTCTTGGTAAAGAGCCGCTAAAGACGTTAAACCGGACGAGCT 1663
Oy 454 GlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIle 473
Db 1662 GTAGCAATCGCGTGGTGGTTCAGGGTGGTCTTCGACGCGAGCTAAACACGCTACTG 1603
Oy 474 LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu----- 487
Db 1602 CTGCTGAGCTTACCCCGCTCTGCTGATGCAACCCGCGGCTGTGATGACGACG 1543
Oy 488 IleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp 507
Db 1542 CTGATCGCGGAAAAACACCACTATCCCGACCAAGCAGCAGAGTGTCTCTACCGCTGAA 1483
Oy 508 AsparGlyValLeuAlaIleAsn---HisAspAspGlyValAsnArgSerIleLeuGlyArg 526
Db 1482 GACAACCACTGCTGCGGTAAACCTCCAT-----GTGCTGACGAGTGACGTTAAACGT 1432
Oy 527 IleSerAspAlaValSerAlaVal-----AlaArgAlaIleLeuProAsp 541
Db 1431 CCGCGTGATTAACAATCTGCTGATACACTGGCAGCATACAGCCGGTGC----- 1384
Oy 542 GluSerGluAsn---GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560
Db 1383 ---ACGGAATATGTTCACTTAATTGATACACCACAGAGGCGCTTATATACGACCCCT 1327
Oy 561 ThrProAlaAspValTyrGlnSerLysLysValPro-----LeuTyrValPheValAla 578
Db 1326 CACCGTCGCCAACGCCCACTAACAATCAACGCCCAACACTTCCACTGGGTGCGCGGA 1267
Oy 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrProGlySerAspThrGlyThrArg 598
Db 1266 CGAGACCTAACCCGAACACCACCGCGCTGGACGAGCACCAGAAATTATGAGCGGC 1207
Oy 599 -----LeuValThrLysPheGlnHis----- 605
Db 1206 CATGACACTATGCTCGCAAGGGATGTTGACGCGCTCCATACATGTTCCGGCTTCCGA 1147
Oy 606 -----AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSer 622
Db 1146 GAGAACCAGATCTCTGATCTGCTACTGCTAGAAAGTTTCCCTGGCTGTAAAGC 1087
Oy 623 ---GluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1086 GTTTCAGTCCGACCGGGCTA-----CCCGCTGCCGATTGTGTAATCCGG 1042
Oy 642 AspGlnLeuArgAlaThr-----LeuGly 649

```

```

Db 1041 AGCAGTTGGTTGGACAGGCGCTGGTATCATCGACGCTGCTGATACCTTCATCTGCAT 982
Oy 650 TyrGlnGlnGluValPheGly-----HisSerThrAsnGlyPheAspLeuSer 665
Db 981 TTCAAAACGCGCGCGCTGGCAATCATGTCGTATTCGCCGAACGCGATGACGTTACT 922
Oy 666 ThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsn----- 680
Db 921 GCC-----TTCCGAACACCCCGCACATCACAGC 895
Oy 681 ---GlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699
Db 894 CAAGGGGCGCGAGCGATGTCATTAACCTGCGTGGAGTGTCCAGC----- 850
Oy 700 AlaProGluThrThrProGlnAspLeuProValAspPheValAsnGlyLysProSerGln 719
Db 849 -----CACTTTCAGAGGTGAATATTCACAT 823
Oy 720 GluAlaLeuLeu-----AlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 822 ACCACGATGCTGTTTATCTCGTGGGATGATGATTAACCGCAGCGGTTCTCGTGGCGCTG 763
Oy 738 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValAlaLysSerSerGlyLeuVal 757
Db 762 ATGCCAAGCTGGGGCAGCTCGCAACGCTACTCTATGACACTATCCAAACAGCGCTGGGCT 703
Oy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 702 TCAGATGCTGATTTTCCTTCGAGCCGACAGCTGCGATCCGACACGCTGACAT 643
Oy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIle 797
Db 642 CGC-----CATCGTTTGTTAACAGCGCGCAGCGCTGGCTGGATT 604
Oy 798 TyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGln 817
Db 603 GAACCGGATGTTTCGACAAAGTACCGCGCGAGATCGGTTTCTCGCGGGGCGACCGC 544
Oy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
Db 543 AGTATTCGTGGCTACAAATATCTATCGTCCGAATATCCCAACGGGTGACGTAAG 484
Oy 838 GlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeu 857
Db 483 GGGGCTCGAAGTTGATTAACCGGATGCTGGAATACCAACCAACGTGACCGGAAATGCG 424
Oy 858 ArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThr 877
Db 423 TGGGGCGCGCTGTTTGTGCTAGTGGCGAAAGCGGTAAAGCATTTGCGCGCAGCAGCTTT 364
Oy 878 LysIleGlyAlaGlyValAlaArgTyrPala-SerProValGlyGlnValArgValAs 897
Db 363 AAAACCGGTACCGCGTGGCGGTGCGCTGGAGATGGCGGTCGCGGCAATCAACTCGCA 304
Oy 897 PValAlaThrGlyVal-----LysGlnGluGlyAsnProIleLysLeuHisPheIle 915
Db 303 TTTTGGCGTACCGCTGCGGATTAAGACGACACGCG-----TTACAGTTTATCAT 253
Oy 915 eGly 916
Db 252 CGGT 249

```

RESULT 8  
 ABR52155  
 ID ABR52155 standard; DNA; 944 BP.  
 AC ABR52155:  
 XX 13-AUG-2002 (first entry)  
 DT  
 XX Haemophilus influenzae strain BASB207 upstream nucleotide sequence.  
 DE  
 XX  
 KW BASB207; antibacterial; anti-inflammatory; vaccine; ss;



```

FT      /tag" a
FT      /product= "N. meningitidis serogroup A antigen"
FT      1..63
FT      sig_peptide
FT      /tag= b
FT      mat_peptide
FT      /tag= c
FT      /product= "Mature N. meningitidis serogroup A antigen"
PN      WO200138350-A2.
XX      31-MAY-2001.
XX      28-NOV-2000; 2000MO-IB01851.
XX      29-NOV-1999; 99GB-0028197.
XX      09-MAR-2000; 2000GB-0005696.
XX      (CHIR-) CHIRON SPA.
XX      (STAT-) STATENS INST FOLKEHELS.
XX      Glut1an1 MM, P1zza M, Rappuoli R, Holst J;
XX      MPI: 2001-381289/40.
XX      P-PSDB: AAU03958.
XX      Novel 85 kDa antigen from Neisseria meningitidis and Neisseria
XX      PT gonorrhoeae, useful in the manufacture of a medicament for treating and
XX      preventing Neisserial bacteria infection -
XX      PS      Claim 6: Flg 2: 92pp: English.
XX
XX      The sequence represents a DNA encoding a Neisseria meningitidis serogroup
XX      CC A 85 kDa antigenic protein. Neisseria meningitidis colonises the pharynx,
XX      CC causing meningitis and, occasionally, septicaemia in the absence of
XX      CC meningitis. This antigenic protein is useful in the manufacture of a
XX      CC medicament for treating or preventing infection due to Neisseria
XX      CC bacteria, such as meningitis and septicaemia. It is also useful as a
XX      CC diagnostic reagent for detecting the presence of Neisseria bacteria or
XX      CC antibodies raised against Neisseria, and as a reagent for raising the
XX      CC antibodies. The Neisserial nucleotide sequences can be expressed in a
XX      CC variety of different expression systems, for example, mammalian cells,
XX      CC baculoviruses, plants, bacteria and yeast.
XX
XX      SO      Sequence 2379 BP; 636 A; 755 C; 596 G; 392 T; 0 other:
XX
XX      Alignment Scores:
XX      Pred. No.: 6.03e-12 Length: 2379
XX      Score: 265.50 Matches: 195
XX      Percent Similarity: 34.28% Conservative: 120
XX      Best Local Similarity: 21.22% Mismatches: 348
XX      Query Match: 5.62% Indels: 257
XX      DB: 22 Gaps: 39
XX
XX      US-09-914-168-2 (1-919) x AAS07278 (1-2379)
XX
XX      QY      116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
XX      DB      46 TCGCCTTTCGATTTGCCCACTTCCACATCCGATCCGTGCGAAGCTTTGGACGCT 105
XX      QY      136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
XX      DB      106 ACCGAGCCGAGCAGCGTATTCAACTACCTGCGGTCGCAAGTGGGAGCACTACAAAC--- 162
XX      QY      150 GluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
XX      DB      163 -----GACACACACGCGGAGTCGATCATCAAAAGCGCTGTAC 198
XX      QY      170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
XX      DB      199 GCCACGGCTTCTTTCGACGACGTACGACGAAACGTCGCGAGGCGGACCTCTGCTGAC 258
XX      QY      187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerLysGlnLysThr 206
XX      DB      1174 TTGGGCTACTTGCACAACGTCAG-----TTTGAT 1203

```

```

DB      259 GTTATGCAAGCCGCCACCATCGGCTCGCTCAACATCACCGGCGCCAAATGCTGCANAAC 318
QY      207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
DB      319 GACGCGC-----ATCAAGAAAAMAACTCGAATCGTTGGGCTGGGCGAGTCGCAATAC 369
QY      225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
DB      370 TTTAATCAGCGACACTCAACACGAGCGAGTCCGGCTGAAAGAAAGAAATACCTCGGGCGT 429
QY      237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
DB      430 GGCAAACTCAATATCCAAATCACGCCCCAAAGTAACCAACTCGCGCCGCAACCGGCTGCAC 489
QY      250 IleAspLeuSerIleIleArgAsnSerIleGlyValAlaAspValIleIleHisAspLeu 269
DB      490 ATCGACATCAGATGACGAGGCGCAATCCGCCAAATCAACCGCATCGAATTTGAA--- 546
QY      270 GlyGluProValTyrIleAspTyrArgAlaValGlnValArgGlyGluGlyAlaAspAsp 289
DB      547 GGCAACCAAGCTATTCGACCGCCGCAACATCGATCGG-----CAG 585
QY      290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
DB      586 ATGTCCTGACACGAGGCGCGCATTTGGACATGGCTGCACAGCAAGCGACCGTTGACCGCC 645
QY      310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPhe 329
DB      646 CAGAAATTCGCCCAAGACATGAAAAGTAACCGACTTCTACCAAGAACAGCGCTACTTTC 705
QY      330 AspGlyArgThrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
DB      706 GATTTCGCTATCCTCGATACCGACATCCAAACCAAGCAACCAAGACGAGACCATC 765
QY      350 SerLeuIleLysThrGlnGlyThrGlnTyrArgPheAspGluValAlaPheIleThrIle 369
DB      766 AAATACCGCTCCACGAAAGGCGGAGCTTCCGCTGGCGCAAGG-----TCGATT 816
QY      370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
DB      817 GAAGGGGACACCAACGAA-----GTCCCAAGCGCGCA 849
QY      390 LeuGlnGlnLeuLeuThrValAlaAsnMetGlyAlaAlaTyrAsnLeuGlnAlaValArgAla 409
DB      850 CTGGAATAACTGCTGACCATGAAAGCCCGCAATGTATGCAAGCCGACGATGACCGCC 909
QY      410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
DB      910 GTTTGGGT----- 918
QY      430 ProGluArgGlnGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArg 449
DB      919 -----GAGATTCGAACCGCATGGCTCGCGCAGCATGCGATACAC 960
QY      450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
DB      961 -----GAAATCAGCGTACAGCGCTGCGCGAAGCGGAGAAC 999
QY      470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
DB      1000 ACCGCGATTTCTCTCGCATCGAATCGGCGGCAAAATTCAGTCAACGAAATTCAC 1059
QY      482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501
DB      1060 ATCAGCGGCAACAAACAAACCGCGAGCA-----GTGTCGGCGCGCAATTCGCCCA 1113
QY      502 LeuTyrAspMetProAspAsp-----ArgValIleValAla 512
DB      1114 ATGGAATTCGCGGCTTACGACACCTCCAAAGCTCAACGCTCAAGAGCGCGTGCAGCTT 1173
QY      513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
DB      1174 TTGGGCTACTTGCACAACGTCAG-----TTTGAT 1203

```

```

QY 533 AlAValAlaArGAlAlLeLeuProAspLusSerGluAsnGluValIleAspLeuProGlu 552
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1204 GCGCTCCCGCTGGCGGACCGCCGACAA-----GTGATTTG----- 1242
QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1243 AAcATGAGCCCTGACCGAAGCTCCAC----- 1269
QY 573 LeuYrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1270 -----GGCTCGCTCGACTTGAGCGCGGC 1293
QY 591 TrpLysSerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610
    |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1294 TGGTTTCAGATACCGG-----TTGGTCATGTCGGCGCGCGGTTCGACGACAAACCTG 1347
QY 611 AspGlyTyrGlnAlaGlyAlaGlnLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1348 TTCGGTACGGGCAAGTCGGCGCCCTCGCCCTCGCAAGCAAA-----ACCAAGCTC 1401
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
    :||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1402 AACGGCTCGCTGCTGTTTACCGACCCGTACTTCACGGCAGACGGGGCTGAGCCTGGGCTAC 1461
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
    :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1462 -----GATATTACGGAAGAACCTTCGACCCGCGCAAGACATGACGACGCTCAACAA 1515
QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArg 690
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1516 TATTAACCCACCCACCGC-----GGCGCGCGGTAGATGGG-TATCCCGCT 1562
QY 691 TyrArgLeu-AspLysLeuLysThrGlnAlaProProGlu-----ThrTyr 705
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1563 TACCGAATACGACCGCGTCAATTTCGGGCTGGCGCGGCAACCTGACCGTCAACACCTA 1622
QY 705 PglAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerG1 719
    |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1623 CAACAAGACACCAACCGCTATGCCGACTTATCAGAAATAGCGCAAAACGACGAGCGC 1682
QY 719 ngluAlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733
    :||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1683 AGACGGCAGCTTCAAGGCTCTGTACAAAGCCGCTGGCGCGCGCAACAGAC 1742
QY 733 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgLysSerLeuGluValGlySe 753
    |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1743 CGACAGCGCGTCAATGCGGACGCGGCTAC----- 1773
QY 753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773
    :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1774 -----CTGACCGCGCTAAATGCGGAATGCCCTGCGCGCGACAAACCTCATA 1823
QY 773 rSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
    :||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1824 C-----TACTCGCCACCCCAACCAACCAACCTGCTCTCC 1859
QY 791 -----GlyIleGlnAlaGlyTyrI1 797
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1860 CTTAAGCAAAACCTTACCGCTGATGCTGGCGGCAAGCTGCGATGGCGGCGCTAC-- 1917
QY 797 eTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyIleAsp1 817
    :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1918 -----GGCAGAACCAAGAAATCCCTTCTTGAACCTTCAAGCGGCGGCGCTGG 1970
QY 817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuTh 837
    |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1971 TTCGGTGGCGGCTACGAAACCGGACGCTGCGCGC-----AAAGTGTATAGACA 2021
QY 837 rGlyGlyGlnValLeuAlaValGly-----ThrAlaGluTyrAsnTyrGluPhe 854
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2022 ATACGGGGAAGAAATACGCTACGGGCGGCAACAAACCAACCTCTCCGCGCGCTGCT 2081

```

```

QY 854 t-----LysAspLeuArgLeuAlaValPheGlyAsp11 865
    :||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2082 CTTCCCGATGCCCGGCGCAAGACGACGACACCGCTCCGCTAGCTGTTTCCACACG 2141
QY 865 eGly----- 866
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 2142 AGCAGCGCTGTGGACGCGAAGCTATACCGCGCGGCAACGTAACAAATCGGT 2201
QY 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyA 883
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2202 TTACTCGGAAGACCGCATTAATCCACTTTACCAAGAAATGCGCTATTCCCGCGCG 2261
QY 883 lGlyValArgTyrPalaSerProValGlyGlnValArgValAspValAlaThrGlyVally 903
    |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2262 CGCGGTACTCTGCTCTGCTTTGGGTCCGATGAATTCAGTACGCTACCGCGTGA 2321
QY 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
    :||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 2322 GAAAAAACCGGAAGACGAATCCACGCTTCCATTCAGCTCGCAGCAGCGTTC 2376
RESULT 10
AAH42129
ID AAH42129 standard; DNA; 2379 BP.
AC AAH42129;
DX 17-SEP-2001 (first entry)
XX
DE Nucleotide sequence of a Neisseria gonorrhoeae protein.
XX
KW Serogroup B protein; outer membrane protein; Neisserial infection;
KM vaccine; ss.
XX
OS Neisseria gonorrhoeae.
XX
FH Key 1. Location/Qualifiers
FT CDS 1..2379
FT FT /*tag= a
FT FT /product= "Neisseria serogroup B protein"
FT sig_peptide 1..63
FT FT /*tag= b
FT FT mat_peptide 64..2376
FT FT /*tag= c
XX
PN WO200152885-A1.
XX
PD 26-JUL-2001.
XX
PF 17-JAN-2001; 2001WO-1B00166.
XX
PR 17-JAN-2000; 2000GB-0001067.
XX
PR 09-MAR-2000; 2000GB-0005699.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Pizza M, Rappuoli R, Giuliani M;
XX
DR WPI: 2001-451895/48.
XX
DR P-PSDB: AAB84745.
XX
PT Composition for treating or preventing infection to, detecting, or for
PT raising antibodies against Neisseria bacteria, comprises an N.
PT meningitidis serogroup B outer membrane preparation and an immunogenic
PT component -
XX
PS Disclosure: Page 64-65; 83pp; English.
XX
CC The present sequence encodes a Neisseria gonorrhoeae protein. The
CC protein is used to produce the compositions of the invention. The
CC specification describes a composition, comprising a Neisseria
CC meningitidis serogroup B outer membrane preparation and an immunogenic
CC component. The immunogenic component is protein disclosed in WO99/57280,
CC WO99/36544, WO99/24578, WO99/66791, WO97/28273, WO96/29412, WO95/03413,

```

CC W099/21132, W099/58683, W099/55873, and/or N. meningitidis protein PorA, CC TdpA, TdpB, PilC, Opa, or Omp85. The composition is used for making CC a medicament for treating or preventing infection due to Neisserial CC bacteria; a diagnostic reagent for detecting the presence of Neisserial CC bacteria or of antibodies raised against Neisserial bacteria; and/or CC a reagent which can raise antibodies against Neisserial bacteria. It may CC also be used as a vaccine.

XX Sequence 2379 BP; 636 A; 755 C; 596 G; 392 T; 0 other;

## Alignment Scores:

Pred. No.:	6 03e-12	Length:	2379
Score:	265.50	Matches:	195
Percent Similarity:	34.28%	Conservative:	120
Best Local Similarity:	21.22%	Mismatches:	348
Query Match:	5.62%	Indels:	257
DB:	22	Gaps:	39

US-09-914-168-2 (1-919) x AAH42129 (1-2379)

```
QY 116 ThrProLeuSerLeuGluGluIleuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
DB   ::::::::::::::::::::
DB 46 TCGCCTTTGGCATTTGGCCACTTCCACATCCGATCCGTCGAAAGCCTTCAGCGT 105
QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
DB   ::::::::::::::::::::
DB 106 ACCGAGCCGACGACCGATTTCACCTACCTGCGCTCAAGTCGGCGACCTTACAC-- 162
QY 150 GluValValProProThrLeuGluProGluLysProGluLeuIleLysArgLeuTyr 169
DB   ::::::::::::::::::::
DB 163 -----GACACACACGCGACGTGCGATCATCAAAAGCCTGTAC 198
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuValLys 186
DB   ::::::::::::::::::::
DB 199 GCCACCGGTTTCTTTCGACGACGAGTCGAAACCTCGGACGGCGAGCTTCTGCTGACC 258
QY 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
DB   ::::::::::::::::::::
DB 259 GTATATGCAACCGCCACCATCGGCTCGCTCAACATCAGCGCGCAAAATGCTCCAAAC 318
QY 207 GluProTyrAlaAsnIleLysAlaIleGluLysP-----IleThrGlnGluSerAla 224
DB   ::::::::::::::::::::
DB 319 GAGCGCC-----ATCAAGAAACCTCGAATCTTCGGCTGCGCGACGTCCCAATAC 369
QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuLysGln----- 236
DB   ::::::::::::::::::::
DB 370 TTTAATTCAGCGCACACTCAACAGGACGTGCGCGCTGAAGAGAAATACCTCGGCGGT 429
QY 237 -----ThrAlaLeuValAlaIleAlaArgAlaValGlyTyrTyrAsp 249
DB   ::::::::::::::::::::
DB 430 GCGCAATCATATTCCAATACACGCCCAAGTAACCAACTCGCGCGCAACCGCGTGCAC 489
QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
DB   ::::::::::::::::::::
DB 490 ATTCGACATCAGCATTCAGCAGGCGCAATCCGCCCAATATCCGACATTCGAATTGAA-- 546
QY 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgIleGluGlyAlaAspAsp 289
DB   ::::::::::::::::::::
DB 547 GCGCAACCAATCTATTCGACCGCCCAATGATGCGG-----CAG 585
QY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyLysPValPheHisHis 309
DB   ::::::::::::::::::::
DB 586 ATGTCCGTCAGCGAAGCGCGCATTTGACATGCGTACACGAAAGCAGCGGTTTCGACCGC 645
QY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
DB   ::::::::::::::::::::
DB 646 CAGAAATATTCGCCAAGACATGCAAAATACCGCACTTCTACAGAAACGCGCTACTTC 705
QY 330 AspGlyArgTyrPheLysAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
DB   ::::::::::::::::::::
DB 706 GATTTTCCGTAATCTCGATACCGACATCCAAACCAAGCAAAACCAAGCAGACATTC 765
QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIle 369
```

```
DB 766 AAATACACCTCCACGAAAGCGGACGTTTCGCTGGCGCAAGTG-----TCGATT 816
QY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
DB   ::::::::::::::::::::
DB 817 GAAGCGCAGCCCAACGAA-----GTCCCAAGGCCCA 849
QY 390 LeuGlnGlnLeuLeuThrValAsnMetGlyAlaTyrAsnLeuGlnAlaValArgAla 409
DB   ::::::::::::::::::::
DB 850 CTGGAAATCTGCTGACCATGAGCGCGCAATGCTCGAAGCCAGACAGATGACCGCC 909
QY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
DB   ::::::::::::::::::::
DB 910 GTTTTCGGT----- 918
QY 430 ProGluArgGluGlnIleGlnAsnAspGluValSerPheGlnGlnSerSerSerArg 449
DB   ::::::::::::::::::::
DB 919 -----GAGATTCAGAACCCGATGGGCTGCGAGCGCTACGCTACACG----- 960
QY 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
DB   ::::::::::::::::::::
DB 961 -----GAAATCAGCGTACAGCGCTGCGCAACCGCGCAACCGCAACCA 999
QY 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
DB   ::::::::::::::::::::
DB 1000 ACCGTCGATTTCGTCGTCGACATCGAACCGCGCGCGGCAAAATCTACGTCACGAAATCCAC 1059
QY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLysAlaArgHis 501
DB   ::::::::::::::::::::
DB 1060 ATCACCAGCAACCAAAACCGCGCAGAA-----GTCTGCGCGCGCGAATTCGCGCA 1113
QY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
DB   ::::::::::::::::::::
DB 1114 ATGGAATCCGCGGCTTACGACACCTCCACAGCTGCAAGCTCCAAAGCGCGCTGACTT 1173
QY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
DB   ::::::::::::::::::::
DB 1174 TTGGGCTACTTCGACCAACGTAACG-----TTTGAT 1203
QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
DB   ::::::::::::::::::::
DB 1204 GCGCTCCCGGCTTCCGCGGTACCGCCGCAAA-----GTGCAATTG----- 1242
QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
DB   ::::::::::::::::::::
DB 1243 AACATGAGCTTACCGCAACCTCCACG----- 1269
QY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
DB   ::::::::::::::::::::
DB 1270 -----GGCTCGCTCGCACTTGAGCGCGGCG 1293
QY 591 TyrGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610
DB   ::::::::::::::::::::
DB 1294 TGGGTTCCAGATATACCGG-----TTGGTATGTCGCGCGCGCTATGCGAGCAACCTG 1347
QY 611 AspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
DB   ::::::::::::::::::::
DB 1348 TTGCGTACGGGCAAGTCGGCGCGCTGCGCGCTCGGAAGCAAA-----ACCAGGCTC 1401
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
DB   ::::::::::::::::::::
DB 1402 AACGCTCGCTGCTGTTTACCGACCGGTACTTCACGCAAGCGGGGTGAGCCTGCGCTAC 1461
QY 651 GlnGlnGluValPheGlnHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
DB   ::::::::::::::::::::
DB 1462 -----GATATTACGGAAGAGCCTTCGACCGCGCAAGCATCGACAGCTCAACAA 1515
QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPAsnArgThrTyrSerLeuArg 690
DB   ::::::::::::::::::::
DB 1516 TATTAACCAACCAACCGCGC-----GGCGGCGCGTAAAGATGGG-TATCCCGCT 1562
QY 691 TyrArgLeu-AspLysLeuLysThrGlnAlaProProGlu-----ThrTr 705
DB   ::::::::::::::::::::
```



```
Db 258 GGCACCGGTTCTTTGACGACGTACGATCGAAATCGCGACGGCTGCTTGTCTGACCC 317
178 -----AsnLysValProArgLeuLysAlaLysPheTyGlnSer 190
Db 318 GTTATGCTATGCCCTACCATCGGCTGCTCACATCACCGGGCCCAAAATGCTCGAGAAC 377
191 SerGln---SerGlyLThrSerAlaIleGlySerSerHisGlnLysThrGluProTyr 209
378 GACCGCATCAGAAAAAACCCTCGAATGCTTCGGGCTGGCGCATCGCAATCTTAATCAG 437
210 AlaAsnIleLysAlaIleLeuGlnAspIleThrGlnGlnSerLametsAspLeuAsnGly 229
438 GCGACACTCACCGAGGCACTCGCGGCTGAAAGAAATAT-----CTCGGGCGC 488
230 SerLleProArgLeuArgLlnThrAlaLeuValAlaIleArgAlaValGlyTyrTyrAsp 249
489 GCGAAACTCAATATTCCAATACAGCCCAAGTACCAAACTCGCCCGCAACCGCGTCGAC 548
250 LleAspLeuSerLleIleArgAsnSerIleGlyGlnValAspValIleIleHisAspLeu 269
549 ATCGACATCACGATTCGACAGCGCAAAATCCGCCAAATCACCGACATCGAATTTGAA--- 605
270 GlyGluProValTyrLleAspTyrArgAlaValGlnValArgGlyGlnGlyAlaAspAsp 289
606 GGCACCAACTCTATTCCGACCGCAACTGATCGG-----CAG 644
290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
645 ATGTGCTGACCGAAGGGCGCATTTGGACATGGCTGACACGAAAGCAGCGGTCGACCGC 704
310 GlyLysTyrGlnThrLysLysAsnLeuIleGlnAsnAlaSerAlaGlnHisGlyTyrPhe 329
705 CAGAAATTCGCCCAAGACATGCAAAAGTACCGACTTACAGCAAGACGGCTACTTC 764
330 AspGlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
765 GATTTCCGATCTCTGATACCGACATCCCAACCAAGCAAAACCGACGACCATC 824
350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIle 369
825 AAATATCAGCTCCAGCAAGCGGACGCTTCGCTGGCGCAAACTG-----TCGATT 875
370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGlnLeu 389
876 GAAGCGACCAACGAA-----GTCCCAAGCGCGGAA 908
390 LeuGlnGlnLeuLeuThrValAsnMetGlyGlnAlaTyrAsnLeuGlnAlaValAlaArgAla 409
909 CTGGAATAAATCTGCTGACCATGAAGCCCGCAATGCTACGAACGCCAGATGACCGCC 968
410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPhe 429
969 GTTTTCGCT----- 977
430 ProGluArgGlnGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArg 449
978 -----GAGATTGAGAAACCCCATGGCTCGCAGCGCTACCATACAC----- 1019
450 ThrGluProAlaGlnValAspGlnSerThrLeuGlnProValIleGlnThrValGlnLeu 469
1020 -----GAAATCAGCGTACAGCCGCTGCCGAACCGCGAACCMAA 1058
470 ThrAspGlyLleLeuMetAspLleSerPro-----IleGln 481
1059 ACCGTCGATTTGCTGTCACATCGAAGCGGAGAAAATCTACGTCMAAGAAATCCAC 1118
482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaIleLysAlaArgHis 501
1119 ATACCGCGCAACAAACAAACCGCGACGAA-----GTCGTGCGCGCGAATTTGCCGCA 1172
502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
::: 111 111
```

---

```
Db 1173 ATGGAATCCGGCCTTACGACACTCCAACTGCAACGCTCCAAAGACGCGTGACCTT 1232
513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
1233 TTGGGCTACTTCGACAAACGTACAG-----TTTGAT 1262
533 AlaValAlaArgAlaIleLeuProAspGlnSerGlnAsnGlnValIleAspLeuProGln 552
1263 GCCGTCGCCGCTCCGCTGATCGCCGCAAA-----GTGCAATTG----- 1301
553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
1302 AACATGAGCTTGACCGAAGCTTCACC----- 1328
573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
1329 -----GGTCGCGTCGACTGACGCGCGC 1352
591 TyrGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610
1353 TTGGGTCAGGATACCGCG-----TTGGTCATGTCGCCCGCGCTATGCGACGACAACTCG 1406
611 AspGlyTyrGlnAlaGlyAlaGlnLeuArgLeuSerGlnAspLysGlyValLysLeu 630
1407 TTCGGTACGGGCAAGTCGCGCGCGCGCTCGCGCAAGCAAA-----ACCAGCCTC 1460
631 TyrAlaThrLysProLeuSerHisProLeuAsnAspLlnLeuArgAlaThrLeuGlyTyr 650
1461 AACGGCTGCGTGCATTACGACCGCTACTTACAGCGCAACGCGGCTGACCTGGGCTAC 1520
651 GlnGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGln 670
1521 -----GATTTTACGGAAGAGCTTCGACCGCGCAAACTGACACGCTCAACMAA 1574
671 HisGlnLleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThrTyrSerLeuArg 690
1575 TATMAAACCAACCAACCGC-----GGCGGCGGCTAAGATGGG-TATCCCGCT 1621
691 TyrArgLeu-AspLysLeuLysThrGlnAlaProGln-----ThrTyr 705
1622 TACCGAATACAGCGCGGTCAATTCGGGCTGGCGGCGGACACCTGACCTCAACACCTA 1681
705 PheLlnAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerGln 719
1682 CAACAAAGCACCAACAGCTATGCCGCTTATCAAACTACGCGCAAAACGACGCGCGC 1741
719 nglnAlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733
1742 AGACGGGACGTTCAAAAGCGCTGCTATCAAAAGCACACTGTGCGGCGGCGCAACAGAC 1801
733 aAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySe 753
1802 CGACAGCGGCTTATGGCGGACGCGCGCTAC----- 1832
753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLleSerGlyValTyr 773
1833 -----CTGACCGCGCTTAATGCGGAATGCGGCTGCCGCGCAAACTGCAATA 1882
773 rSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
1883 C-----TACTCCGCCACCCACACCAACCAACTGGTTCTTCCC 1918
791 -----GlyLleGlnAlaGlyTyrIle 797
1919 CTTAAGCAAAACCTTACCGCTGATGCTCGCGGCGAAGGTGGCATTCGGGCGGCTAC-- 1976
797 eTTPSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAspGln 817
1977 -----GGCAGAACCAAAAGAAATCCCTCTTTGAAAACTTACAGCGCGCGGCTGGG 2029
817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProLleSerAspLysGlyTyrLeuTh 837
2030 TTCGGTGGCGGCTTACGAAAGCGGACGCTGCGCGC-----AAAGTATATGACGA 2080
```





Db	679	GA	CTTCTACCAAAATACCGGCTACTTCGATTTCCGATATCTTCGATACCGCATCTCAACAAC	738
Qy	341	Ile	LeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg	360
Db	739	AACGAAGACAAACCAAGCAGACCATCAAAATCATCCGTCACGAAGCGGACGTTCCCGT	798	
Qy	361	Pha	AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro	380
Db	799	TGGGGCAAGTC	-----TCCATCGAAGCGCACACCAAGAA-----	834
Qy	381	Asp	LysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGlu	400
Db	835	-----GTCCCCAAAGCGAAGTGGAAAAACATGCTGACCATGAAGACCGCGCAAA	882	
Qy	401	Ala	TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe	420
Db	883	TG	TACGAACGCCAGCATGACCGCGTTTGGGT-----	918
Qy	421	Asn	MetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal	440
Db	919	-----	-----GAGATTCAGAACCGCATGGGC	939
Qy	441	Ser	PheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu	460
Db	940	TCGGCAGGCTACGCATACAGC	-----GAAATCAGCGTA	972
Qy	461	Glu	ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--	479
Db	973	CAGCGCTCCGAACGCTAAACCAACCGTCGATTTTCGTCCTGCATCGAACCGGGC	1032	
Qy	480	-----	-----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu	492
Db	1033	CGAAAAATCTACGTCACACGCAATACATACACCGCAACAAACACCGCGACGAA--	1089	
Qy	493	Asn	LeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp--	508
Db	1090	-----GTGCTCGCGGTGAATTCGCCAAATGGAAATCCGCACCTTACGACACCTCCAAGCTG	1146	
Qy	509	-----	-----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle	523
Db	1147	CAAGCTTCAACAGCGCGCTCGAGCTTTTGGGCTACTTCGACATGTCCAG--	1197	
Qy	524	Leu	GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer	543
Db	1198	-----	-----TTTGATGCTGTCCCGCTTGCAGCGCACCGCCGACAAA--	1233
Qy	544	Glu	AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla	563
Db	1234	-----GTGATTTG	-----AACATGATGCTCGACCGAACGTTCCACC--	1269
Qy	564	Asp	ValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg	583
Db	1269	-----	-----	1269
Qy	584	Asp	GlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr	601
Db	1270	--GGTTCCTCGATTGACGCGCGGTTGGGTTCAGATACCGG--	-----TTGGTCA	1320
Qy	602	Lys	PheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu	621
Db	1321	TCGCGCAGGGGTTCCCAACACAACTGTTCGCTACGGGCAAGTCGGCGCAGCTCGCGCC	1380	
Qy	622	Ser	CluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn	641
Db	1381	TCGAGGAGCAAA-----ACCAGCTTAACGGCTCGCTGTCGTTACTACCGCTACTTC	1434	
Qy	642	Asp	GlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly	661
Db	1435	ACGGCAGCGGGTCAGCTGGGCTAC-----GATGTTTACGAAAGAGCTTCGACCGC	1488	
Qy	662	Phe	AspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly	681

Db	1489	CCCAAGCAGTTCGACCAGCATCAACAATATATAAACACCACCGCA	-----GGC	1539
Qy	682	GlyTyrAsnArg	-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu	696
Db	1537	GCAGGCATCCGATGAGCGTGCCTGTTACCGAATACGACCGCTGAATTTTCGCTTGGTG		1596
Qy	697	LysThrGlnAlaProGluThrTrpGlnAspLeuPro	-----ValAspPheVal	713
Db	1597	GCAGAACACCTGACCGTCAACACCTACAACAAGCCGCCAACACACTATGCCGACTTTATC		1656
Qy	714	Asn	-----ProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLys	730
Db	1657	AAGAAATACGGCAAAACCGCAGCAGCGCAGCTTCAAAGCGCTGGCTGTACAANGT		1716
Qy	731	ThrVal	-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg	744
Db	1717	ACCGTCGGCTGGGGCGCAACAAACCGACAGCGCTTATGGCCGACGCGCGCTAC		1773
Qy	745	GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle		764
Db	1774	-----	-----CTGACGGCGCTGAACGCGCAANATC	1797
Qy	765	AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg		784
Db	1798	GCCTCCCTGGCAGCAAACTGCAATAC	-----TACTCCGCC	1833
Qy	785	AlaHisGlnMetThrGly	-----	790
Db	1834	ACCCACAAACCAACTGGTCTTCCCCCTCGACAAACCTTCCAGCTGATGCTCGCGCGC		1893
Qy	791	-----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg		808
Db	1894	GAAGTCGGCATTTGCGGCGGCTAC	-----GGCAGAACCAAGCAAAATCCCTCTCTTT	1944
Qy	809	LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer		828
Db	1945	GAAGAACTTCTACGCGCGCGCTGGTTCGGTCGCGGATACCAAGACGCGCACGCTCGT		2004
Qy	829	ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly	-----	845
Db	2005	CCG	-----AAAGTCTATGACGAATACGGCGAAAAATACGCTACGCGCGCAACAAA	2055
Qy	846	ThrAlaGluTyrAsnTyrGluPheMet	-----LysAsp	856
Db	2056	AAAGCCACAGTCTCCGCCAGCGTCTTCCCGATGCCCGCGCGAAGACGCGCGCAC		2115
Qy	857	LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp	-----	870
Db	2116	GTCCGCGCTGAGCTTGTTCGCCAGCAGCAGCGCTGTGGGACGCCAAACCTACGACGAC		2175
Qy	870	-----	-----	870
Db	2176	AAACAGCAGTTCGCGCACCGCGCGCAGGGTTCAAAACATTTACGGCGCGGCAATACCCAT		2235
Qy	871	---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSer		889
Db	2236	AAATCCACCTTTTACCAACAGAAATGGCGCTATTCGCGCGCGCGCGGTACTCTGCTCTCG		2295
Qy	890	ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu	-----GluGlyAsn	907
Db	2296	CCTTTAGGCCCATGAAATTCAGTACGCTACCCCGCTGAAGAAAAAACCGGAAGACGAA		2355
Qy	908	ProIleLysLeuHisPhePheIleGlyThrProPhe	919	
Db	2356	ATCCAACGCTTCCAATTCCTCAACTTCGCGCAGCGATTC		2391

RESULT 13

RESOL 13  
AAS07279  
ID AAS07279 standard: DNA: 2394 BP.

XX AAS07279:

XX  
AC  
AA301213;

DT 23-OCT-2001 (first entry)



```
Db 1033 CGGAAATCTACGTCAAGAAATCCATCACCGCGCAACAAACACCGCGAGGAA --- 1089
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp- --- 508
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1090 ---GTCGGTGGCCGCGAATCGCGCAATGGAATCCCGCGCTTACGACACCTCCAGAGTG 1146
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 509 -----ArgValLeuAlaAlaIleAsnHisAspAspGlyValAlaAsnArgSerIle 523
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1147 CAACGCTCCAAAGAGCGGTGAGCTTTGGGTACTTCGACACAGTACAG --- 1197
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1198 -----TTTGATCGCTCGCGCTTGGCGGCGACACCGCAACAA --- 1233
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1234 -----GTCGATTTG-----AACATGAGCCTTGACCGAGACGTTCACC --- 1269
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1269 ----- 1269
Qy 584 AspGlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr 601
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1270 ---GGCTCGCTCGACTTGAGCGCGGCTGGGTACAGGATACCGG- ---CTGGTCA TG 1320
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1321 TCCGAGCGGTTTCCAGACAACTGTCGTGACGGCAAGTCGGCGCGCTGCGCGCC 1380
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1381 TCACGAAGCAAA-----ACCAGCGTCAACGGCTCGCTGTTTACCGACCGCTACTTC 1434
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1435 ACGCAGACGGGTGACGCTGGGTAC-----GATGTTTACGGAAAGCCTTCGACCCG 1488
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1489 GCCTAACGATCGACCGATCAACAAATATAAACACCCACCGCA-----GGC 1536
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1537 GCAGGCATCCGATGAGCGTGCTGTTACCGAATACGACCGCGTGAATTTTCGGTTTGGTG 1596
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1597 GCAGAACACCTGACCGTCAACACCTTACAAACAAAGCCCAACACATATGCCGACTTTATC 1656
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1657 AAGAAATACGCAAAACCGACGCGCACAGCGCATTCCAAAGGCTGCTGTACAAAGGT 1716
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1717 ACCGTGCGCTGGGGCGCAACAAACCGACAGCGGTTATGGCCGACGCGCGCTAC--- 1773
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetalIle 764
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1774 -----CTGACGGGCGTGAACCCCGAAATC 1797
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1798 GCCCTGCCGCGCAGCAAACTGCAATAC-----TACTCCGCC 1833
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 785 AlaHisGlnMetThrGly----- 790
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1834 ACCCAACAAACCAACCTGTTCTTCCCTTAAGCAAAACCTTCACGCTGATGTCGGCGGC 1893
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
```

```
Db 1894 GAAGTCGCATTTGGCGGGGTACT-----GCAGAACCAAAAGAAATCCCTTCTTT 1944
Qy 809 LeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 1945 GAAAACTTCTACGGCGCGGCTGGGTTCGGTCCGGGATACGAAAGCGGCACGCTCGGT 2004
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValIcIy----- 845
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 2005 CCG-----AAAGTGTATGACCAATACGGCGAAAAAATCAGCTACGGCGGCAACAA 2055
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 2056 AAAGCAACGCTCCGCGAGCTGCTTCCCGATGCCCGCGCGCAAGACGCGGCACC 2115
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 2116 GTCCGCTGAGCCTGTTTCCCGACCGCAGCGCGTGTGGGACGGCAAAACCTACGACGAC 2175
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 870 ----- 870
Db 2176 AACAGCAGTTCCGCGACGCGCGAGGTTTCAAAACATTTACGGCGCGGCAATACCCAT 2235
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 2236 AAATCCACCTTTACCAACGAATTCGCTATTCGCCGCGCGCGGTACCTGGCTCTCG 2295
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 2296 CTTTATAGCCGATGCAATTCAGCTACGCTACCGCTTCAAGAAAAACCGGAAGACGAA 2355
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 2356 ATCCAACGCTTCCAATTCCAACTCGCAGCAGCTTC 2391
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
RESULT 14
AAH42128
ID AAH42128 standard; DNA; 2394 BP.
AC AAH42128;
DT 17-SEP-2001 (first entry)
DE Nucleotide sequence of a Neisseria serogroup B protein.
KW Serogroup B protein; outer membrane protein; Neisseria infection;
KW vaccine; ss.
OS Neisseria meningitidis.
FH Key Location/Qualifiers
FT CDS 1..2394
FT /tag= a
FT /product= "Neisseria serogroup B protein"
FT sig_peptide 1..63
FT /tag= b
FT mat_peptide 64..2391
FT /tag= c
PN WO200152885-A1.
XX 26-JUL-2001.
XX 17-JAN-2001; 2001WO-IB00166.
XX 17-JAN-2000; 2000GB-0001067.
XX 09-MAR-2000; 2000GB-0005699.
XX (CHIR-) CHIRON SPA.
XX Pizza M, Rappuoli R, Giuliani M;
XX WPI; 2001-451895/48.
XX P-PSDB; AAB84744.
DR
```

Db 547 GGCAACCAAGTCTATTCCGACCGCAAACTGATGCGG-----582

```
QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1381 TCCAGGACAAA-----ACCACGCTTAACGCGTCGCTGTCTTACTACCGCTACTTC 1434
QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
Db 1435 ACGCAGACGGGTCAGCGTGGGTAC-----CATGTTTACGGAANAAGCCTTCGACCCG 1488
QY 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 1489 CGCAAGCATCGACGACGATCAAAATATAAACCACGCGCA-----GGC 1536
QY 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
Db 1537 GCAGGCATCGCATGAGCGTCGCTTACCGCAATACGACGCGTGAATTTTCGTTGGTG 1596
QY 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
Db 1597 GCAGACACCTGACCGTCACACCTACACAAAGCGCCCAACACTATGCGCAGCTTATC 1656
QY 714 Asn-----GlyLys-----ProSerGlnGlnAlaLeuAlaGlyValAlaValHisLys 730
Db 1657 AAGAAATACGCAAAACCGCAGCGCAGCAGCTTCAAAAGCGTGGCTGTACAAAGGT 1716
QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 1717 ACCGTCGGTGGGGCGCACAACACCGCAGCGCTATGCGCGCAGCGCGGTAC--- 1773
QY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 1774 -----CTGACGGCGGTGAACGCCGAATC 1797
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 1798 GCCTCGCTGGCAGCAAACTGCAATAC-----TACTCGGCC 1833
QY 785 AlaHisGlnMetThrGly----- 790
Db 1834 ACCCACAACCAACCTGGTTCTTCCCTGAGCAAAACCTTCACGCTGATGCTCGGCGGC 1893
QY 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808
Db 1894 GAAGTCGGCATTCGGGCGCGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944
QY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
Db 1945 GAAACTTCTACGCGCGCGCGCTGGGTTCGGTGGCGGATACGAAACGCGCACGCTCGGT 2004
QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
Db 2005 CCG-----AAAGTCTATGAGATACGGCGAAATAATCAGCTACGGCGGCAACAA 2055
QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 2056 AAAGCCAACGCTCCGCGGAGCTGCTTCCCGATCCCGCGCGGAAAGACGCGCGCAC 2115
QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 2116 GTCCGCTGAGCTGTGTTGCCGAGCGAGCGAGCGGTGTGGGACGCGCAAAACCTACGAC 2175
QY 870 ----- 870
Db 2176 AACAGCAGTTCGCGACCGCGCGGTTCAAAACATTTACGGCGCGCGCAATACCAT 2235
QY 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
Db 2236 AAATCCACCTTTTACCAACGAATTCGGCTATTCCCGCGCGCGCGGTTCACCTGCTCG 2295
QY 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 2296 CTTTAGCCCGATGAATTCAGCTACGCTACCGCTGAAGAAAAAACCGGAAGACGAA 2355
QY 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
```

```
Db 2356 ATCCAACGCTTCCAAATTCCAACTCGGCACGAGCTTC 2391
```

## RESULT 15

```
AAH42130
ID AAH42130 standard; DNA; 2394 BP.
```

```
XX AC AAH42130;
```

```
XX XX 17-SEP-2001 (first entry)
```

```
XX DE Nucleotide sequence of a Neisseria serogroup A protein.
```

```
XX XX Serogroup A protein; outer membrane protein; Neisserial infection;
KW vaccine; ss.
```

```
XX OS Neisseria meningitidis.
```

```
XX FH Key Location/Qualifiers
FT CDS 1..2394
```

```
FT FT /*tag= a
```

```
FT FT /product= "Neisseria serogroup A protein"
```

```
FT FT 1..63
```

```
FT FT /*tag= b
```

```
FT FT 64..2391
```

```
FT FT /*tag= c
```

```
XX PN WO200152885-A1.
```

```
XX XX 26-JUL-2001.
```

```
XX XX 17-JAN-2001; 2001WO-IB00166.
```

```
XX XX 17-JAN-2000; 2000GB-0001067.
```

```
XX PR 09-MAR-2000; 2000GB-0005699.
```

```
XX XX (CHIR-) CHIRON SPA.
```

```
XX PI Pizza M, Rappuoli R, Giuliani M;
```

```
XX XX WPI; 2001-451895/48.
```

```
XX DR P-PSDB; AAB84746.
```

```
XX XX Composition for treating or preventing infection to, detecting, or for
```

```
XX PT raising antibodies against Neisserial bacteria, comprises an N.
```

```
XX PT meningitidis serogroup B outer membrane preparation and an immunogenic
```

```
XX PT component .
```

```
XX PS Disclosure; Page 70-71; 83pp; English.
```

```
XX CC The present sequence encodes a Neisseria serogroup A protein. The
XX CC protein is used to produce the compositions of the invention. The
XX CC specification describes a composition, comprising a Neisseria
XX CC meningitidis serogroup B outer membrane preparation and an immunogenic
XX CC component. The immunogenic component is protein disclosed in WO99/57280,
XX CC WO99/36344, WO99/24578, WO99/66791, WO97/28273, WO96/29412, WO95/03413,
XX CC WO99/31132, WO99/58683, WO99/55873, and/or N. meningitidis protein PorA,
XX CC TbpA, TbpB, PilC, Opa, or Omp85. The composition is used for making
XX CC a medicament for treating or preventing infection due to Neisserial
XX CC bacteria; a diagnostic reagent for detecting the presence of Neisserial
XX CC bacteria or of antibodies raised against Neisserial bacteria; and/or
XX CC a reagent which can raise antibodies against Neisserial bacteria. It may
XX CC also be used as a vaccine.
```

```
XX SQ Sequence 2394 BP; 655 A; 758 C; 586 G; 395 T; 0 other;
```

## Alignment Scores:

```
Pred. No.: 2,06e-11 Length: 2394
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: 22 Gaps: 40
```

US-09-914-168-2 (1-919) x AAH42130 (1-2394)

Qy 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135  
Db 46 TCGGCTTTGGCAGCTTGGCGACTTCCACCATCAAGAGTCCGCGTGAAGGTTGGCAGGT 105  
Qy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149  
Db 106 ACCGAGCGGAGTACCGTATTCAACTACCTGCCGCTCAAGTCCGGGACACCTACAAAC--- 162  
Qy 150 GluValValProProThrLeuGluProGluLysProGlyLeuIleAspArgLeuTyr 169  
Db 163 -----GACACACGCGAGTGCATCATCAAAAGCCTGTAC 198  
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186  
Db 199 GCCACCGGTTCTTTGACGAGTACGCGCTCGAAGTCCGGGACGGGACGCTCTGCTGACC 258  
Qy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206  
Db 259 GTTATCGAAGCCGCCACCATCGCTCGCTCAACATCACCGCGCAAAATGCTGCAGAAAC 318  
Qy 207 GluProTyrAlaAsnIleLysAlaLeuGluAsp-----IleThrGlnGluSerAla 224  
Db 319 GACGCC-----ATTAAAGAAACCTCGAATCGTTCCGGCTGGCGCAGTCGCAATAC 369  
Qy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236  
Db 370 TTTAATCAGGCGACACTCAATCAGGAGTCCGCGGCTCGAAGAGAGATACCTCGGGGCG 429  
Qy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249  
Db 430 GCGAAACTCAATATCCAAATCAGCCCAAGTAACCAAACTCGCCCGCAACCGCTCGAC 489  
Qy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269  
Db 490 ATCGACATCAGCATTTGACGAGGGCAAAATCCGCCAAATACCGACATCGAATTTGAA--- 546  
Qy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluAlaAspAsp 289  
Db 547 GCGAACCAAGTCTATTCGACCGGCAAACTGATGCGG----- 582  
Qy 280 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306  
Db 583 -----CAGATGTGCTGACCGAAGGCGCATTTGGACATGG 618  
Qy 307 -----PheHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320  
Db 619 CTGACACGAAGCAACCAATTCACGAGCAGAGAAATTTGCCAAGACATGGAAAGTAACC 678  
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspVal 340  
Db 679 GACTTCTACGAGAACAGCGGTCTTCCGATTTCCGATCTCCGATACCGACATCCAAACC 738  
Qy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360  
Db 739 AACGAAGCAAAACCAAGCAGACCATCAAAATCACCGTCCACGAAGGCGGAGTTCGGT 798  
Qy 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380  
Db 799 TGGGCAAAAGTC-----TCCATCGAAGGCGCACCAACGAA----- 834  
Qy 381 AspLysLeuProValLysArgGluLeuGluGlnLeuLeuThrValAsnMetGlyGlu 400  
Db 835 -----GTCCCAAGCCCACTGGAAATCTGCTACCATGAAGCCCGGCAAA 882  
Qy 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420  
Db 883 TGGTACGAACGCCACGATGACCGCGCTTTTGGGT----- 918  
Qy 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440  
Db -----

Db 919 -----GAGATTTCAGAACCGCATGGCG 939  
Qy 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460  
Db 940 TCGGACAGCTACGCATACAGC-----GAAATCAGCGTA 972  
Qy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLysSerPro--- 479  
Db 973 CAGCGCTGCCAAACCGCGAAACCAACCGTCGTGATTCCTGCATCATCAACCGGCG 1032  
Qy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492  
Db 1033 CGAAATTCACGTCAACCAAAATCCACATCCCGCAACCAACCAACCGCACGAA--- 1089  
Qy 493 AsnLeuValAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508  
Db 1090 ---GTCGTGCGCGGAATTGGCCAAATGNAATCCGGCCCTTACGACACCTCCAAGCTG 1146  
Qy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523  
Db 1147 CAACGCTCCAAAGAGCGGTTCGAGCTTTTGGGCTACTTCGACAACGTACAG----- 1197  
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543  
Db 1198 -----TTTGATGCGCTCCCGCTTCGCGGCACACCCGACAAA--- 1233  
Qy 544 GluAsnGluValIleAspLeuProGluThrAlaLeuAlaAsnArgLysThrProAla 563  
Db 1234 -----GTCGATTTG-----AACATGAGCCTGACCGAAGCTTCACC----- 1269  
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
Db 1269 ----- 1269  
Qy 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601  
Db 1270 ---GGCTCGCTCGACTTGAGCGGGCTGGGTACAGGATACCGGC-----CTGGTCAATG 1320  
Qy 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1321 TCCGAGGCGTTTCCCAAGACAACTGTTCGTACGGCAAGTCGCGCCCTCGCGGCC 1380  
Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
Db 1381 TCACGAAGCAAA-----ACCAGCTCAACGCTCGCTCGCTTTACCGCACCGTACTTC 1434  
Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661  
Db 1435 ACGGACAGCGGGTCAGCTGGGCTAC-----GATGTTTACGGAAGAGCTTCGACCGG 1488  
Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681  
Db 1489 CCGAAGCATCCACGACATCAACAATAATAAACACCAACCGCGCA-----GGC 1536  
Qy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696  
Db 1537 GCAGGCATCCGATGAGCGTGTGTACCGAATACGACCGCGTGAATTCGGTTGGTG 1596  
Qy 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713  
Db 1597 GCAGAACACCTGACCGTCAACACCTACAAACGCGCCAAACACATATGCCACTTTATC 1656  
Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730  
Db 1657 AGAATACGGCAAAACCGCACGCGCAGCGAGCTTCAAAGGCTGGCTGTACAAAGGT 1716  
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgIleTyrArg 744  
Db 1717 ACCGTGCGGTGGGGCGCAACAAACCGACAGCGCTTATGGCGCGACGCGGTAC--- 1773  
Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764  
Db 1774 -----CTGACGGCGGTGAACGCCGAATC 1797

```
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
DB 1798 GCCCTGCCGCGCAGCAAACTGCAATAC-----TACTCGGCC 1833
QY 785 AlaHisGlnMetThrGly----- 790
DB 1834 ACCCACACCAACCACTGGTCTCCCTTAAAGCAAACTTCACGCTGATGCTCGGCGGC 1893
QY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
DB 1894 GAAGTCGGCATTCGGCGCGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944
QY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
DB 1945 GAAAACTTCACGCGCGCGCTGGTTCGGTGGCGGATACGAAAGCGGCACGCTCGGT 2004
QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
DB 2005 CCG-----AAAGTGTATGAGCAATACGGCGAAATAATCAGCTACGGCGGCAACAAA 2055
QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
DB 2056 AAGCCCAACGTCTCCGCGAGCTCTCTCCGATGCCGCGGCGAAAGCGCGCACCC 2115
QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
DB 2116 GTCGCGCTGACGCTGTTCGCGAGCGAGCGGTGTGGAGCGGCAAAACCTAGACGAC 2175
QY 870 ----- 870
DB 2176 AACAGCAGTTCGCGCGCGCGGCGGTTCAAAACATTTACGGCGCGGCAATACCAT 2235
QY 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTtrPalaser 889
DB 2236 AAATCCACCTTTACCAACGAATTGCGCTATTCCGCGCGCGCGGTACTGCTGCTCG 2295
QY 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
DB 2296 CTTTAGCGCGGATGAATTCAGCTACGCTACCGCTGAGAAACCAACCGAAGACGAA 2355
QY 908 ProIleLysLeuHisPheIleGlyThrProPhe 919
DB 2356 ATCCAACGCTTCCAATTCCAACTCCGCGACGACGCTC 2391
RESULT 16
AAA81478
ID AAA81478 standard; DNA; 52253 BP.
AC AAA81478;
XX
DB 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_26 SEQ ID NO:26.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN W0200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US233573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
```

```
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
XX Claim 7: Page 532-547; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
```

Sequence 52253 BP; 12877 A; 15510 C; 13103 G; 10761 T; 2 other;

#### Alignment Scores:

Pred. No.:	1.17e-09	Length:	52253
Score:	259.00	Matches:	191
Percent Similarity:	33.48%	Conservative:	121
Best Local Similarity:	20.49%	Mismatches:	342
Query Match:	5.48%	Indels:	278
DB:	21	Gaps:	40

US-09-914-168-2 (1-919) x AAA81478 (1-52253)

```
QY 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
DB 47343 TCGCCTTTGGCAGCTTGGCGAGCTTCACCATCCAGACATCCCGCTCGAAGGCTTCAGCGT 47402
QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
DB 47403 ACCGAGCGGAGTACCGGTATTCACTACTGCGCGTCAAAAGTCGGCGACACCTCAAC--- 47459
QY 150 GluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
DB 47460 -----GACACACACGCGAGTGCATCATCAAAAGCCTGTAC 47495
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
DB 47496 GCCACCGGTTTCTTTGTGACGAGCTACCGGTCCGAAACTCCGCGAGCGAGCTCTCTGACCC 47555
QY 187 PheTyrGlnSerSerGlnSerGlnGlyThrSerAlaIleGlySerSerHisGlnLysThr 206
DB 47556 GTTATCGAACGCCCCACCATCGGCTCGCTCAACATCACCGCGCGCAAAAANTCTCGAANAAC 47615
QY 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
DB 47616 GACGCC-----ATTAAGAAAAACCTCGAATCGTTCGGGCTGGCGCAGCTCGCAATAC 47666
```



```
QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 47667 TTTAATCAGCGCACACTCAATCAGCGAGTCGCGCGCTGAAAGAAAGTAATACCTCGCGCGC 47726
QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrArgPhe 249
Db 47727 GCGAAACTCAATATCCAAATCAGCGCAAGTAACCAACTCGCCGCAACCGCGTCGAC 47786
QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyValAlaAspValIleIleHisAspLeu 269
Db 47787 ATCGACATCAGCATTCAGCAGGCGCAATCCGCCAAATCAGCAGCATCGAATTGAA--- 47843
QY 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
Db 47844 GCGAACCAAGTCTATTCGCCAGCGCAAACTGATCGCG----- 47879
QY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 47880 -----CNAATGTCCTGACCGCAAGCGCGCAITTTGGACATGG 47915
QY 307 -----PheHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
Db 47916 CTGACACGAAGCAACCAATTCACGAGCAGAAATTTGCCCAAGATATGAAAGTAACC 47975
QY 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340
Db 47976 GACTTCTACCAAAATAACGGCTACTTCGATTTCCTGATCCGATACCGCAGCATCCAAACC 48035
QY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360
Db 48036 ACGAAGCAAAACCAAGCAGACCATCAAAATCACCGTCCAGAGCGGAGGTTCCTGT 48095
QY 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 48096 TGGGGCAAGTC-----TCCATCGAAGCGCACCAACGAA----- 48131
QY 381 AspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGlu 400
Db 48132 -----GTCCCCAAACCGCAACTGGAAAACTGCTGACCATGAAGCCCGCAAA 48179
QY 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 48180 TGGTACGAACGCCAGCATGACCGCGTTTGGGT----- 48215
QY 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
Db 48216 -----GAGATTCAAGAACCGCATGGGC 48236
QY 441 SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 48237 TCGGAGCGGTACGCATACAGC-----GAAATCAGCGTA 48269
QY 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
Db 48270 CAGCGCGTCCCGAACGCTGAACCAAAACCGTCGATTTCGTCTGCATCGAACCGGCG 48329
QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
Db 48330 CGGAAATCTAGCTCAACGAAATACATCACCGGCAACAAACAAACCGCGCAGCAA--- 48386
QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
Db 48387 ---GTCGTCCCGCGTGAATTCAGCAAAATGGAAATCCGCACTTACGACCTCCCAAGCTG 48443
QY 509 -----ArgValLeuAlaIleAsnHisAspGlyValAlaAsnArgSerIle 523
Db 48444 CAACGTTCCAAAGCGGCTCAGCTTTTGGGCTACTTCGACAATGTCCAG----- 48494
QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
Db 48495 -----TTTGATGCTGCTCCCGCTTGGCGGCGCAGCGCCGACAA--- 48530
QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 48531 -----GTCGATTGG-----AACATGAGTCTGACCGAAGCTTCCACC----- 48566
QY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 48566 ----- 48566
QY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 48567 ---GGTTCCCTGGATTTCAGCGCGGTTGGGTTCAAGATACCGG-----TTGGTTCATG 48617
QY 602 LysPheGluHisAsnLeuLeuAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 48618 TCCGCAAGGCGTTTCCCAAGACACACCTGTTGCTGACGGCAAGTCGCGCGCAGCTCGCGCC 48677
QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 48678 TCCAGAGCAAA-----ACCAGCTTAACGGCTCCGCTGCTGTTTACTGACCGCTACTTC 48731
QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSerThrAsnGly 661
Db 48732 ACGGCAGACGGGTGCTGAGCTGGCTAC-----GATGTTTACGGAAGGCTTTCGACCCG 48785
QY 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 48786 CGCAAGACATCGACCATCAACCAATATAAAACCAACCCACCGCA-----GGC 48833
QY 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
Db 48834 GCAGGCATCCGATGAGCTGCTGTTCCGAATACAGCCGCTGATTTTCGTTGGTTGGTG 48893
QY 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
Db 48894 GCAGAACACCTGACCGTCAACACCTACACAAAGCGCCCAACACATATGCGCATTTATC 48953
QY 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
Db 48954 AAGAAATACGGCAAAACCGCACGACGCGAGCTTCAAAAGCTGGGTGTACAAAGGT 49013
QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 49014 ACCGTGGGTGGGGCGCAACAAACCGACAGCGCGTTATGGCCGCGCGCTAC--- 49070
QY 745 GlnArgTyrSerLeuGluValGlySerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 49071 -----CTGACGGCGTGAACCGCGAAATC 49094
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 49095 GCCCTGCTGGCAGCAAACTGCAATAC-----TACTCCGCC 49130
QY 785 AlaHisGlnMetThrGly----- 790
Db 49131 ACCCAACCAACCAACCTGTTCTTCCCTCGACCAAAACCTTACGCTGATGCTCGCGGC 49190
QY 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808
Db 49191 GAAGTCGGCATTCGGGCGCGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 49241
QY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
Db 49242 GAAACCTTCTACGGCGCGGCTGGGTTCGGTCCGCGGATACGAAAGCGCGCAGCTCGGT 49301
QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
Db 49302 CCG-----AAAGTCTATGACGAATACGGCGGAAAAAATCAGCTACGGCGCGCAACAA 49352
QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 49353 AAAGCAAGCTCTCCGCGCGAGCTGCTTCCCGATCCCGCGCGGCGGAGCGCGCAC 49412
QY 857 LeuArgLeuAlaPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 870 ----- 870
```

```
Db 49413 GTCCGCTGACCCGTTTGGCCGACGACGAGCGTGTGGGAGCGGCAAAACCTACGACGAC 49472
Qy 870 ----- 870
Db 49473 AACAGCAGTTCGCGACGCGCGGAGGTTTCAAAACATTTTACGCGCGGCAATACCCAT 49532
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
Db 49533 AAATCCACCTTTACCAAGCAATTCGCTATTCCGCGCGCGGCTTACCTGCTCTCG 49592
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 49593 CCTTAGCCCGCATGAATTCAGTCAGTCAGCTACCGCTGAAGAAAAACCGAGACAGAA 49652
Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 49653 ATCCAAGCGTTCCAATTCCAACTCGGACGACGTTTC 49688

RESULT 17
AAF21544/c
ID AAF21544 standard; DNA; 349980 BP.
XX
AC AAF21544;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Massignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
XX WPI; 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
PT
XX
PS Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
```

```
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 other.;

Alignment Scores:
Pred. No.: 1.41e-08 Length: 349980
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: Gaps: 40

US-09-914-168-2 (1-919) x AAF21544 (1-349980)
Qy 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
Db 180761 TCGCCTTTGGCACTTGGCGACTTCACCATCCAAGACATCCGCGTCCAAGGCTTCGAGCGT 180702
Qy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
Db 180701 ACCGACCGGAGTACGCTATTCAACTACCTGCCCGTCAAGTCGGCGACACCTTACAAAC--- 180645
Qy 150 GluValValProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 180644 -----GACACACAGCGCAGTGCATCATCAAAAGCCCTGTAC 180609
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
Db 180608 GCCACCGGTTTCTTTGACGAGCTAGCGCTCGAAACTGCGGAGCGGCGAGCTCTGCTGACC 180549
Qy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 180548 GTTATCGAAACGCCGCCACCATCGGCTCGCTCAACATCACCGCGGCAAAATATGCTGCAAAAC 180489
Qy 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
Db 180488 GAGGCC-----ATTAAGAAAAACCTCGAATCTGTCGGGCTGGCGCAGTCGCAATAC 180438
Qy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 180437 TTTAATACGCGCACACTCAATCAGCAGTCGCGCGGCTGAAAGAAAGATACCTCGGCGCC 180378
Qy 237 -----ThrAlaLeuValAlaAlaArgAlaValAlaGlyTyrTyrAsp 249
Db 180377 GGCAAACTCAATATCCAATCAGCCCAAGTAACCAACTCGCCCGCAACCGGCTCGAC 180318
Qy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
Db 180317 ATCGACATCAGCATTTGACGAGGGCAAAATCCGCCAAATCACCAGCATCGAATTTGAA--- 180261
Qy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
Db 180260 GGCAACCAAGTCTATTCCGACGCGCAAACTGATGCGG----- 180225
Qy 290 LysAlaPheThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 180224 -----CAAATGTCTTCGACCGAGGCGGATTTGGACATCG 180189
Qy 307 -----PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
Db 180188 CTGACACGAACCAACCAATTCACGACGAGCAAAATTTGCCCAAGATATGAAAAAATAACC 180129
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340
Db 180128 GACTTCTACCAAAAATAACGGGTACTTCGATTTCCGTATTCCTCGATACCGACATCCAAACC 180069
Qy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360
```

D

Bb 179270 GCAGGCATCCGCGATGACGGCTGCCTGTATTACGAATACACGCCGTGAATTTTCGGTTTGGTG 179211

Qy 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713  
|||||  
Dbb 179210 GCAGAACAACCTGACCGGTCAACACTACAAACAAAGCGCCCAACACACTATGCCGACTTTATC 179151

Qy 714 Asn-----GlyLys---ProSerGlnGluIuaLeuLeuAlaGlyValalavalHisLys 730  
|||||  
Dbb 179150 AAGAAATACGGCAAAAACCGACGACAGCGCAGCTTCAAAGCTGGGTGTCAAAAGT 179091

Qy 731 ThrVal-----AlaAspAsnLeuValasnProMetArgGlyTyraG 744  
|||||  
Dbb 179090 ACCGTGCGCTGGGGCCAAACAAACCGACAGCGCTTATGGCCGACGCGCGCTAC--- 179034

Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValserAspAlasMetAlalle 764  
|||||  
Dbb 179033 -----CTCAGGGCGCTGAACGCCGCAAAATC 179010

Qy 765 AlaArgAlaGlylleSerGlyValTyrSerPheGlyAspAsnAlatyrGlySerAsnAaG 784  
|||||  
Dbb 179009 GCCCTGCCTGGCAGCAAACTGCAATAC-----TACTCCGCC 178974

Qy 785 AlahisGlnMethrGly----- 790  
|||||  
Dbb 178973 ACCCACAAACAAACCTGGTTCTTCCCCTGAGCAAAACCTTCACGCTGATGCTCGCGCGC 178914

Qy 791 -----GlylleGlnAlaGlyTyrilleTrpSerAspAsnPheAsnHisValProTyrArG 808  
|||||  
Dbb 178913 GAAGTCGCATTCGGGGCGCTAC-----GGCAGAACCAAGAAATCCCTCTCTTT 178863

Qy 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlahisAspSerLeuSer 828  
|||||  
Dbb 178862 GAAAACTTCTACGGCGCGCCCTGGTTCGGTCGCGGATACGAAAGCGCACGCTCGGT 178803

Qy 829 ProfileSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845  
|||||  
Dbb 178802 CCG-----AAAGTCTATGACGAATACGGCGCAAAATCAGCTACGCGCGCACACAAA 178752

Qy 846 ThrAlaGluTyrAsnTyrGluPhMet-----LysAsp 856  
|||||  
Dbb 178751 AAAGSCCAACGCTCCGGCGAGCTCTTCCCGATCCCGCGCGAAAGACGCGCGCAC 178692

Qy 857 LeuArgLeuAlaValPheGlyAspLleGlyAsnAlatyrAsp----- 870  
|||||  
Dbb 178691 GTCCGCTGACCTGTTCGCCGACGAGCGAGCTGTGGGACGCGCAAAACCTACGACGAC 178632

Qy 870 ----- 870  
|||||  
Dbb 178631 AACAGCATTTCCGCGACC CGCGCAGGGTTCAAAACATTTACGGCCCCGGCAATACCCAAT 178572

Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValcglyValArgTrpAlaser 889  
|||||  
Dbb 178571 AAATPCACCTTACCAACGAATTCGCTATTCCCGCGCGCGCGGTTCACCTGGCTCTCG 178512

Qy 890 ProValGlyGlnValArgValaspValAlaThrGlyValLysclu-----GluGlyAsn 907  
|||||  
Dbb 178511 CCTTAGCCCGATGAATTCACCTACGCTACCCGCTGAAGAAAAAACCGGAAGACGAA 178452

Qy 908 ProfileLysLeuHisPhePheIleGlyThrProPhe 919  
|||||  
Dbb 178451 ATCCAACGCTTCCAATTCCAACTCGCAGCAGCTTC 178416

RESULT 18  
AAAL1516  
ID AAAL1516 standard; DNA; 2394 BP.  
XX  
XX AAAL1516,  
XX  
XX  
DT 21-AUG-2000 (first entry)  
XX  
XX DNA encoding outer membrane protein (omp) 85.  
XX  
KW Outer membrane protein; omp85; gonococcal infection; symptomatic disease

KW meningococcal infection; protective immune response; vaccine; ss.

```

XX Os  Neisseria meningitidis.
XX FH  Location/Qualifiers
XX Key  1..2394
XX CDS  /*tag= a
        /product= "outer membrane protein 85"
XX
XX WO200023595-A1.
XX
XX PD  27-APR-2000.
XX
XX XX  22-OCT-1998; 98WO-US22352.
XX
XX XX  22-OCT-1998; 98WO-US22352.
XX
XX PA  (UIMO-) UNIV MONTANA.
XX
XX PI  Judd RC, Manning SD;
XX DR  WPI: 2000-339694/29.
XX DR  P-PSDB; AAY84947.
XX
XX XX  New isolated outer membrane protein 85 of Neisseria gonorrhoeae and N.
XX PT  meningitidis useful for vaccine, therapeutic and diagnostic
XX PT  compositions for gonococcal or meningococcal infections -
XX XX
XX PS  Claim 46; Page 85-89; 98pp; English.
XX
XX CC  The present sequence encodes an outer membrane protein (omp) 85 of
XX CC  Neisseria meningitidis. The omp polypeptides and polynucleotides are
XX CC  useful in compositions for use in the prevention, treatment and diagnosis
XX CC  of non-symptomatic gonococcal infection or meningococcal infection and
XX CC  symptomatic disease. They are also useful for the detection of
XX CC  hybridisation complexes. Antigens and antibodies specific omp proteins
XX CC  also provide diagnostic, therapeutic and prophylactic compositions for
XX CC  the treatment or prevention of the infections described above. The
XX CC  antibodies are useful for inducing a protective immune response in
XX CC  humans or animals with N. gonorrhoeae, N. meningitidis, or other
XX CC  Neisseria species. The proteins, antibodies and polynucleotide
XX CC  sequences of the present invention may also be used in the screening
XX CC  and development of chemical compounds such as drugs or vaccines.
XX
XX SQ  Sequence 2394 BP; 656 A; 752 C; 587 G; 399 T; 0 other;

Alignment Scores:
Pred. No.: 6.35e-11 Length: 2394
Score: 253.00 Matches: 190
Percent Similarity: 33.37% Conservative: 121
Best Local Similarity: 20.39% Mismatches: 343
Query Match: 5.35% Indels: 278
DB: 21 Gaps: 40

US-09-914-168-2 (1-919) x AAA15156 (1-2394)
Qy 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
   :::::::::::::: ::::
Db 46 TCGCCTTTGGCATTTGGCGACTTCACCATCCAAAGACATCCGCTGTGGAAGGCTTGCGAGCGT 105
   :::::::::::::: ::::
Qy 136 AsnAsp-----TyrIleProGluThrGlnGlyGluGlnProAsnSer 149
   ::::
Db 106 ACCGACGGGACGACCGGATTCACATCTACCTGCGCTGAAAGTCGGCGATACCTACAAC--- 162
   ::::
Qy 150 GluValValProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
   ::::
Db 163 -----GACACACACGGCAGTGCATCATCAAAAGCCTGTAC 198
   ::::
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
   ::::
Db 199 GCCACCGGTTTCTTTGACGAGGTACGGCTGAAACTGCGGACGGGACGCTCCTGCTGACC 258
   ::::
Qy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206

```

```

Db 259 GTTATCGAAGCCGCCACCATCGCTCAACATCACCGCGCAAAATAGCTCGCAAAAC 318
   ::::
Qy 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGluGluSerAla 224
   ::::
Db 319 GACGCC-----ATTAAGAAAAACCTCGAATCGTTGGGCTGGCGCAGTCGCAATAC 369
   ::::
Qy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgLeuGln----- 236
   ::::
Db 370 TTTAATCAGCGGCACACTCAATCAGCGAGTCGCCGCGCTGAAAGAAGAAATACCTCGGGCGC 429
   ::::
Qy 237 -----ThrAlaLeuValAlaAlaArgAlaValAlaGlyTyrTyrAsp 249
   ::::
Db 430 GGCAAACTCAATATCCAAATCAGCCCAAGTAACCAAACTCGCCCGCAACCGCTCGAC 489
   ::::
Qy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAlaValIleIleHisAspLeu 269
   ::::
Db 490 ATCGACATCAGCATTTGACGAGGGCAAAATCCGCCAAAATCACCAGACATCGAATTTGAA--- 546
   ::::
Qy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
   ::::
Db 547 GGCAACCAAGTCTATTTCGACCGCAAACTGATGCGG----- 582
   ::::
Qy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
   ::::
Db 583 -----CAGATGTCGCTGACCGAAGCGCGCATTTGGACATGG 618
   ::::
Qy 307 -----PheHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
   ::::
Db 619 CTGACACCAACCAACCAATTCACGAGCAGAAATTTGCCAAGACATCGAAAAAGTAACC 678
   ::::
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspVal 340
   ::::
Db 679 GACTTCTACCAAGAACACAGCGTACTTCGATTTCCGTATCCTCGATACCGACATCCAAACC 738
   ::::
Qy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrClnTyrArg 360
   ::::
Db 739 AACGAAGACAAACCAACGACGACCATCAAAATCACCGTCCACGAAGCGCGGAGTTTCGT 798
   ::::
Qy 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
   ::::
Db 799 TGGGGCAAGTC-----TCCATCGAAGCGGACACCAACGAA----- 834
   ::::
Qy 381 AspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGlu 400
   ::::
Db 835 -----GTCCCAAGCCGCAACTGGAATAAATCGTACCATGAAGCCCGGCAAA 882
   ::::
Qy 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
   ::::
Db 883 TGGTACGAACCCGACGACATGACCGCGCTTTTGGGT----- 918
   ::::
Qy 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
   ::::
Db 919 -----GAGATTCAACACCGCATGGGC 939
   ::::
Qy 441 SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
   ::::
Db 940 TCGGACGCTACGCAATACAGC-----GAAATCAGCGTA 972
   ::::
Qy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
   ::::
Db 973 CAGCGCTGGCCAAACGCGCAACCAACCGTCGATTTTCGCTCGCATCGAACCGGCGC 1032
   ::::
Qy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
   ::::
Db 1033 CGGAAATCTACGTCAACGAAATCCATCACCGGCAACCAACAAACCCGCAACGAA--- 1089
   ::::
Qy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
   ::::
Db 1090 ---GTCTGTCGCGCGCAATTCGCGCAATTCGCGCAATTCGCGCGCTTACGACACCTCCAAGCTG 1146
   ::::
Qy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
   ::::

```

Db 1147 CAACGCTCCAAAGACGCGTTCAGCTTTTGGCTACTTCGACAACTACAG----- 1197  
Qy 524 LeuGlyArgIleSerAspAlaValIleSerAlaValAlaArgAlaIleLeuProAspGluSer 543  
Db 1198 -----TTTGATGCGCTCCGCTTCGCGGCACACCGACAAA--- 1233  
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563  
Db 1234 -----GTCGATTG-----AACATGAGCGTGACCGAACGTTCCACC----- 1269  
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
Db 1269 ----- 1269  
Qy 584 AspGlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr 601  
Db 1270 ---GGCTCGCTCGACTTGAGCGCGGCTGGGTACAGGATACCGC-----CTGGTCATG 1320  
Qy 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1321 TCCGACGCGCTTCCCAAGACAACCTGTTTCGTACGGGCAAGTCGGCGCCTCGCGGCC 1380  
Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
Db 1381 TCACGAAGCAAA-----ACCACGCTCAACGCGCTCGGTGCTTTACCGACCCGTACTTC 1434  
Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661  
Db 1435 ACGCAGACGGGTGAGCTGGCTAC-----GATGTTTACGGAAGACCTTCGACCGC 1488  
Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681  
Db 1489 CGCAAGACATCGACACGATCAACAATATAAAACCAACACCGGCA-----GGC 1536  
Qy 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696  
Db 1537 CGAGCATCCGATGAGCGTCCCTGTACCGAATACGACCGCGTGAATTCGGTTTGGTG 1596  
Qy 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713  
Db 1597 GCAGAACACCTGACCGTCAACACCTCAACAAGCGGCCAACACACTATGCCGACTTATC 1656  
Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730  
Db 1657 AAGAAATACGGCAAAACCGACGCGCACGACGCGACGTTCAAAGCGCTGCGCTACAAAGGT 1716  
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744  
Db 1717 ACCGTGCGCTGGGCGGCAACAAACCGACAGCGGTATATGGCGCGCGGTAC--- 1773  
Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764  
Db 1774 -----CTGACGGCGTCAACGCGCAAAATC 1797  
Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
Db 1798 GCCCTGCCGCGCAGCAAACTGCAATAC-----TACTCCGCC 1833  
Qy 785 AlaHisGlnMetThrGly----- 790  
Db 1834 ACCCACACCAACCACTGGTTCTTCCCTTAAGCAAAACCTTCACGCTGATGCTCGCGCGC 1893  
Qy 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808  
Db 1894 GAAGTCGCAATGCGGGCGCTAC-----GCACGAACCAAGAAATCCCTTCTTT 1944  
Qy 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLysSer 828  
Db 1945 GAAAACTTCTACGGCGCGCGCTGGGTTCGGTTCGCGGATACGAAAGCGGACGCTCGGT 2004  
Qy 829 ProfileSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845  
Db 2005 CCG-----AAAGTGTATGACGAATACGGGAAAAAATCAGCTACGCGGCGCAACAA 2055

Qy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856  
Db 2056 AAAGCCAACTCTCGCGGAGCTGCTCTCCGATGCTGCGGCGGAAGACGCGCGCAC 2115  
Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870  
Db 2116 GTCGCGCTGAGCCTGTTTTCGCGACGAGCAGCGTGTGGGACGGAACCTTACGACGAC 2175  
Qy 870 ----- 870  
Db 2176 AACAGCAGTTCCGCGACGCGGCGGAGGTTCAAAACATTTACGCGCGCGCAATACCCTAT 2235  
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889  
Db 2236 AAATCCACCTTTACCAACGAATTCGCTATTCCGCGCGCGCGGCTTACCTGGCTCTCG 2295  
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907  
Db 2296 CTTTAGCCCGCATGAATTCAGGTACGCTACCGCTGAAGAAAAAACCGGAAGACGAA 2355  
Qy 908 ProfileLysLeuHisPhePheIleGlyThrProPhe 919  
Db 2356 ATCCAACGCTTCCAATTCCAACTCGGCACGAGCTTC 2391

## RESULT 19

AAI97964  
ID AAI97964 standard; DNA; 6617 BP.  
XX  
AC AAI97964;  
XX  
DT 20-NOV-2001 (first entry)  
XX  
DE Lawsonia intracellularis coding sequence SEQ ID NO: 1.  
XX  
KW HtrA; PonA; HycC; YefW; ABC1; Omp100; Lawsonia intracellularis infection;  
KW vaccine; ds.  
XX  
OS Lawsonia intracellularis.  
XX  
PN JP2001169787-A.  
XX  
PD 26-JUN-2001.  
XX  
PF 20-OCT-2000; 2000JP-0320736.  
XX  
PR 22-OCT-1999; 99US-0160922.  
XX  
PA (PFIZ ) PFIZER PROD INC.  
XX  
DR WPI; 2001-592540/67.  
XX  
PT Lawsonia intracellularis polynucleotide and encoded protein, used to  
PT prevent Lawsonia intracellularis infection -  
PS Claim 3; Page 37-39; 67pp; Japanese.  
XX  
CC The present invention provides isolated polynucleotides encoding HtrA,  
CC PonA, HycC, LysS, YefW, ABC1 or Omp100 protein of Lawsonia  
CC intracellularis. The sequences can be used in vaccines for the prevention  
CC of Lawsonia intracellularis infection. The present sequence is a coding  
CC sequence of the invention.  
XX  
SQ Sequence 6617 BP; 2216 A; 1153 C; 1142 G; 2106 T; 0 other;

## Alignment Scores:

Pred. No.: 3.65e-09 Length: 6617  
Score: 238.50 Matches: 198  
Percent Similarity: 34.72% Conservative: 137  
Best Local Similarity: 20.52% Mismatches: 391  
Query Match: 5.05% Indels: 239  
Gaps: 38  
DB:

US-09-914-168-2 (1-919) x AAI97964 (1-6617)

Qy	50	AsnGlnAlaLysAlaGlyAsnProProValLeuLeuLeuThrProGluClnIleGlnAlaArg	69	 :::     :::	Db	3764	AATGCTGCTTCAAAAGACGATCCTTCTATTTGGTGTCTCCCATTTCAAATTAATGGCTCA	3823	
Qy	70	LeuAsnAlaAlaGlyLeuAsnAlaLys---ProGlnSerGlnAlaLeuAspValValAsn	88	 :::     :::	Db	3824	TCAAATGATGAAGAGTTTCAACACAGAACTACCAATGCTTCTTGCACATGTCATTAAAGAAAT	3883	
Qy	89	PheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAsp	108		Db	3884	-----	-AAGGGA	3889
Qy	109	MetSerValIleGluGluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSer	128	 :::     :::	Db	3890	TTTCTGTGCTCCCTTAATAAATCTGCATTAATCTT-----CTATATAACCAAAATATC	3943	
Qy	129	ThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn	148	 :::     :::	Db	3944	TCCCAACTTAATATTTCTACTGCCAAAAAGGTA-----GCTCAACAACCTCCAT	3991	
Qy	149	SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu	168	 :::     :::	Db	3992	GCTGACTATGTAGTA-----	4006	
Qy	169	TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr	188	 :::     :::	Db	4007	TACGCCAGTTTCATCAACAGCTGAAAT-----TTTAGT	4042	
Qy	189	GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSer-----HisGlnLys	205	 :::     :::	Db	4043	ATTGATAGTAGGCTTATTGATAGTACAGGTGTAGCATCTGCACGTCCCATTATACATAGAA	4102	
Qy	206	ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMet	225	 :::     :::	Db	4103	AAACCAAAATTTAATGAGCTAAATATTGCTGTACACAGAACTTCTGCAAGCTATAAGT	4159	
Qy	226	AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuAlaAlaArgAlaVal	245	 :::     :::	Db	4160	-----AATGGCTTTATAAGAAA---AACACTATTGCTGTATACGTATTATCGGCTT	4210	
Qy	246	GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle	265	 :::     :::	Db	4211	AAAGTTCTTGATCCTGTATCTTACACGACTCATTATTAATAAGGGAGATCATACT	4270	
Qy	266	IleHis-----	274	 :::     :::	Db	4271	GATCATGCCAAATTAATGCAGCAATCAAAAATATGGAATTAGGA-----TATTTT	4324	
Qy	275	IleAspTyrArgAla---ValGluValArgGlyGluGlyAlaAspLysAlaPheThr	293	 :::     :::	Db	4325	AGTGATGCTCTGCAAGTATTCAAGAAAGCGGGAAGGA-----CGATTACTTGTA	4375	
Qy	294	ThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGlu	313	 :::     :::	Db	4376	TTTACTGTACAGAAAGCCCTAAAATACAGATGTGTGTTCAGGCTCAAAAGCTGTA	4435	
Qy	314	ThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyr	333	 :::     :::	Db	4436	AGTATCGATAACATCTTGCTGCAATGAGTTCTTAAAAAAGGA-----	4477	
Qy	334	LeuAspArgSerVal-----AspValIleLeuProAspAsnThrAlaAspValSerLeu	351	 :::     :::	Db	4478	-----TCAGTTATTAGTGATAGACTATTGTCCCAAGATATTCAAAAATATCCCGAC	4528	
Qy	352	IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspPro	371	 :::     :::	Db	4529	CTCTATAGAAAGAGGCTACTACTCTCGCTGAA-----GTTAATATGAATAAAGAG	4582	
Qy	372	LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlu	391	 :::     :::	Db	4583	AAAGAAAATACTTCTCTGCA-----	ACA	4606



```
Db 592 AAACGCGATTTTCCGCTACCGGTCGCGGATAAAGACGACACCGG-----TTACAG 642
QY 913 PhePheIleGly 916
Db 643 TTTTACATCGGA 654

RESULT 21
AAS93087/C
ID AAS93087 standard; cDNA: 1425 BP.
AC AAS93087;
XX
XX
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #28891.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX
XX P-PSDB; ABG28900.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
XX
XX Claim 1: SEQ ID No 28891; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1425 BP; 366 A; 407 C; 353 G; 299 T; 0 other;

Alignment Scores:
Pred. No.: 8.59e-10 Length: 1425
Score: 235.50 Matches: 56
Percent Similarity: 57.268 Conservative: 15
Best Local Similarity: 45.168 Mismatches: 48
```

```
Query Match: 4.98% Indels: 5
DB: 23 Gaps: 2
US-09-914-168-2 (1-919) x AAS93087 (1-1425)

QY 795 GlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGly 814
||||| : : : : : |||||
Db 1134 GCGTGGATTGAACCGGTGATTTCGACAAAGTACCGCGGATCTCGCTTTCTTGGCCGG 1075
||||| : : : : : |||||

QY 815 GlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGly 834
||||| : : : : : |||||
Db 1074 GCGACCGCAGTATTGCGTGCTACAAATACAAATCTATCGTCCGAAATACGCCAACGCT 1015
||||| : : : : : |||||

QY 835 TyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMet 854
||||| : : : : : |||||
Db 1014 GACCTGAAGGGCGCTCGAAGTTGATACCGGATCGCTGGAAATACCAAGTACCGTGACC 955
||||| : : : : : |||||

QY 855 LysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThr 874
||||| : : : : : |||||
Db 954 GGAATATGGTGGCGCGGGTGTTCGATAGTGGCGAAGCGGTAAAGCGATATTCGCCGC 895
||||| : : : : : |||||

QY 875 AsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlnVal 894
||||| : : : : : |||||
Db 894 AGCGACTTTAAACCGGTACCGCGGTGCGCTGGGAATCGCGGTGCGGCCAATC 835
||||| : : : : : |||||

QY 895 ArgValAspValAlaThrGlyVal-----LysGluGluGlyAsnProIleLysLeuHis 912
||||| : : : : : |||||
Db 834 AAACGCGATTTTTCGCTACCGGTGCGCGGATAAAGACGACACACGCGG-----TTACAG 784
||||| : : : : : |||||

QY 913 PhePheIleGly 916
||||| : : : : : |||||
Db 783 TTTTACATCGGA 772

RESULT 22
AAX91724
ID AAX91724 standard; DNA: 2325 BP.
XX
XX AC AAX91724;
XX
XX DT 25-AUG-1999 (first entry)
XX
XX DE Porphyromonas gingivalis protein PG45 encoding DNA.
XX
XX KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
XX KW vaccine; antigenic; ds.
XX
XX OS Porphyromonas gingivalis.
XX
XX PN WO9929870-A1.
XX
XX PD 17-JUN-1999.
XX
XX PF 10-DEC-1998; 98WO-AU01023.
XX
XX PR 04-AUG-1998; 98AU-0005028.
XX PR 10-DEC-1997; 97AU-0000839.
XX PR 31-DEC-1997; 97AU-0001182.
XX PR 30-JAN-1998; 98AU-0001546.
XX PR 10-MAR-1998; 98AU-0002264.
XX PR 09-APR-1998; 98AU-0002911.
XX PR 23-APR-1998; 98AU-0003128.
XX PR 05-MAY-1998; 98AU-0003338.
XX PR 22-MAY-1998; 98AU-0003654.
XX PR 29-JUL-1998; 98AU-0004917.
XX
XX (CSLC-) CSL LTD.
XX
XX PI Agius CT, Barr IC, Hocking DM, Margetts MB, Patterson MA;
XX PI Ross BC, Rothel LJ, Webb EA;
XX
XX WPI: 1999-385613/32.
XX
XX P-PSDB; AAY34506.
XX
```



PT Antigenic Porphyromonas gingivalis peptides for preventing  
PT gingivitis  
XX  
PS  
Claim 12; Page 222-223; 588pp; English.  
XX  
CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic  
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAX91536 to  
CC AAX91801. AAX91802 to AAX91899 represent PCR primers used in the  
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial  
CC activity with a vaccine mechanism of action. The PG polypeptides can be  
CC used as vaccines especially against Porphyromonas gingivalis. Probes can  
CC be used to detect Porphyromonas gingivalis in standard hybridisation  
CC assays. Porphyromonas gingivalis is involved in periodontal disease  
CC especially gingivitis.  
XX  
SQ Sequence 2325 BP; 575 A; 636 C; 547 G; 567 T; 0 other;

Alignment Scores:  
Pred. No.: 5.53e-09 Length: 2325  
Score: 229.00 Matches: 161  
Percent Similarity: 32.99% Conservative: 93  
Best Local Similarity: 20.91% Mismatches: 264  
Query Match: 4.84% Indels: 253  
DB: 20 Gaps: 35

US-09-914-168-2 (1-919) x AAX91724 (1-2325)

QY 300 ProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeuIle 319  
DB 325 CCGGTTTCATCAGTCAGTC-----AAATCCGATAGCGGCTAAGTGGCG 372  
QY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAsp 339  
DB 373 ACGAACATCTCCGCGAACACGGTACTTCGATGCTAAA---GTAAAAAGCAGTGTGACC 429  
QY 340 ValIleLeuProAspAsn---ThrAlaAspValSerLeuIleTyrAspThrGlyThrGln 358  
DB 430 ACTCTGAAAGAGACTCGCTCAAGCGCAAAATCTCTATACGGTGATATGCGCTCTCT 489  
QY 359 TyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThr 378  
DB 490 TATCATAGCAGCATCATTCCTTACCGATC-----AGCAT 528  
QY 379 AspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMet 398  
DB 529 TTCCCGCAGCATCTTCGCTTACAGGCAGCTCCGCT-----TTGATCAGGAAA 579  
QY 399 GlyGluAlaTyrAsnLeu-----GlnAlaValArgAlaLeuSer 411  
DB 580 GGAGACCATTTCAATTTGGCAAAGCTGCACGAAGAGCGTCAGACCATCATGCTCCCTGTG 639  
QY 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431  
DB 640 ACAGAC-----AATGGTTACTACTTCCGCCCCACAGATATATCTACGAAGCC 690  
QY 432 ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGlu 451  
DB 691 GATACCTCTCGTAAAGAGTGCCTGATGCTGCGAGCAAGCTCTCGGAAGATATCTCCA 750  
QY 452 ProAla-----GlnValAspGluSerThrLeu----- 460  
DB 751 CCCCAGGCATCGCCCGTGGAGGATAGGAAACGACAGCATGCTGCTCGGAATGAAC 810  
QY 461 ---GluProValIleGluThrValGluLeuThrAsp---GlyIleLeu----- 474  
DB 811 GGAGAAAGCCACAGACTCGCTCGAAGTGGAGGATATGAAGTCTCTTACTATCGTAAA 870  
QY 475 MetAspIleSerPro-----IleGluPheSerAlaSerAsnLeuIle 488  
DB 871 ATGCGGTTTCGCCCAAGATTTTGGCCAAACGCTTTCGTTTCTTCCGCGCAATCTGTAT 930  
QY 489 GlnAspLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp 508  
DB 1830 CGTGATATAGCTATGCG----- 1848

DB 931 CGGCAGAAA----- 939  
QY 509 ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528  
DB 940 -----GACGATGAGACGACGCAAAATCTTGGCTCGCTTGGGA 978  
QY 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIle 548  
DB 979 ---GCCTTCTCGCTTATCGATCTCAATTTTTCACAGCGGATCTCATTTCCGCGCTTTTG 1035  
QY 549 AspLeuProGluArgThrAlaLeu-----AlaAsnArgLysThrProAla 563  
DB 1036 GATGTGCGACTGCTAAACACCTTCGAAACCTTGGGATGCATCATAGAGAC----- 1089  
QY 564 AspValTyrGlnSerLysLys-----ValProLeuTyrValPheValAla 578  
DB 1090 ---TTGTTTACGAGCAAAAGCAATGACTTCATCGGTCGCGGACTGAATTTGCTTGTCT 1146  
QY 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTrp-----GlySer----- 593  
DB 1147 CGCGCAATGATTCGCGGAGGAGAGAAATCTTCTTGAATATCGTGGATCGTATGAG 1206  
QY 594 ---AspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGly 612  
DB 1207 TGGGAGACCGCAATCGT---CCCGAAATAGCAGCAATCGGCTGATGATATAAATTCG 1263  
QY 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla 632  
DB 1264 TACAACATCAATACGCGCGTGAACCTCTCGTTCCCTCGATGTTATTTCCCGCTGTGCTG 1323  
QY 633 ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThr----- 647  
DB 1324 GATAAATACTATTACTACCCAGACTACGACTTTTTCAGGCTTCTGCCACCGCGCTGAAC 1383  
QY 647 ----- 647  
DB 1384 AGGCGACACTACTTTAGCATGTACTCTTTCGGCTTTTCGACCACTACGAATTTTCAGCCC 1443  
QY 648 -----LeuGlyTyrGlnGlnGluValPheGlyHis 657  
DB 1444 TCCAAGAACACCGCATGCTATTTTCCCGCTCAAGCTCACTACAACTCTCTGGGCGCAT 1503  
QY 658 SerThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHisGluIleSer 674  
DB 1504 CAGACAGAACTTTCAGGCGCATTCAGCGCAACAATCCGCCCCCTGTCTGCTACGCTTCA 1562  
QY 675 ArgSerIle---IleGlnAsnGlyGlyTyrPheAsnArgThrTyrSerLeuArgTyrArg--- 692  
DB 1563 GAGTCAGTTCTTCTGCTCAATGCGGTATATCTATACGTTTCAACAAATCCGCTTTCAGAGAA 1622  
QY 693 -----LeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAsp 707  
DB 1623 AAGTCTCATCATCTTTGGATGCAATTCGACT----- 1655  
QY 708 LeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGly----- 725  
DB 1656 -----ATCCGAGCGCAGCAATCTCTCTGAATCTCATCTATCTGCGAGCGG 1700  
QY 726 -----ValAlaValHis-LysTh 731  
DB 1701 CAAGAAGTACGCGACACCAAGAAATTCGTCGCGCTCCCTCTCTCAGTTCATCAAGC 1760  
QY 731 rValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluVal 751  
DB 1761 CACGGGAGAA-----CTCGCTATTTCTTATACCAT 1790  
QY 751 lGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAla---ArgAlaGlyIleSe 770  
DB 1791 A-----GACCGCAATCAGTCACTGGCAACCGCTTTCGGGACAGG 1829  
QY 770 rGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrG1 790  
DB 1830 CGTGATATAGCTATGCG----- 1848

```
Qy 790 yGlyIleGlnAlaGlyTyrIleTrrpSerAspAsnPheAsnHisValProTyrArgLeuAr 810
Db 1849 -----AATATGCGAGTGGCACCCCTATAGCGAGCA 1877
Qy 810 gPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu----- 827
Db 1878 GTTCATGTAGCGGTGCATATAGTATCAGAGCTTTCACCGTCGCGTACATCGGCCCGCG 1937
Qy 828 -----SerProIleSerAspLysGlyTyrLeu-----ThrGlyGlyGlnValLe 842
Db 1938 ACGGTTCAATCCGGATCCGACAATCAGTATTCCTATTTGGATCAGTGGCGGAATTCAA 1997
Qy 842 uAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPh 862
Db 1998 ACTCGAAGCCCAACGTGGGAATATAGAGCAAGCTTTTCGGGGATCTCCACGAGCGGTTT 2057
Qy 862 eGlyAspIleGlyAsnAlaTyr----- 869
Db 2058 CCTCGATCGCGGCAACGTTTGGCTCTTGAGGAGGATTCCTCCGTCGCGGGGTGCTCT 2117
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValG1 884
Db 2118 GTCCGAAGTGGATCGGTGGAGCAATTCCTCGAATAGCATCGCTCTCGGCACCGGTGTCGG 2177
Qy 884 yValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys-- 903
Db 2178 CCTCGCTACGATCGGCATTTCTCGTGGTTCGTGTGATGTCGGCTTCGGCTCCACCT 2237
Qy 904 -----GluGluGly-----AsnProIleLy 910
Db 2238 TCCTTACATACGGGTAGAAGGTACTACAATATCCACGCTTAAAGGATGCCATCGG 2297
Qy 910 sLeuHisPhePheLeGlyThrProPhe 919
Db 2298 TTTCATTTGGCTGCGGCTATCCCTTC 2325

RESULT 23
AAX91597
XX AC AAX91597;
XX DT 25-AUG-1999 (first entry)
XX DE Porphyromonas gingivalis protein PG45 ORF encoding DNA.
XX KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
XX OS vaccine; antigenic; ds.
XX OS Porphyromonas gingivalis.
XX PN WO9929870-A1.
XX PD 17-JUN-1999.
XX PF 10-DEC-1998; 98WO-AU01023.
XX PR 04-AUG-1998; 98AU-0005028.
XX PR 10-DEC-1997; 97AU-0000839.
XX PR 31-DEC-1997; 97AU-0001182.
XX PR 30-JAN-1998; 98AU-0001546.
XX PR 10-MAR-1998; 98AU-0002264.
XX PR 09-APR-1998; 98AU-0002911.
XX PR 23-APR-1998; 98AU-0003128.
XX PR 05-MAY-1998; 98AU-0003338.
XX PR 22-MAY-1998; 98AU-0003654.
XX PR 29-JUL-1998; 98AU-0004917.
XX (CSLC-) CSL LTD.
XX PA Agilus CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
XX PI Ross BC, RotheL LJ, Webb EA;
```

```
XX WPI; 1999-385613/32.
DR P-PSDB; AAY34379.
XX Antigenic Porphyromonas gingivalis peptides for preventing
PT gingivitis
XX Claim 12; Page 136-137; 588pp; English.
XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis.
XX Sequence 2409 BP; 603 A; 657 C; 569 G; 580 T; 0 other:
SO Alignment Scores:
Pred. No.: 5.79e-09 Length: 2409
Score: 229.00 Matches: 161
Percent Similarity: 32.99% Conservative: 93
Best Local Similarity: 20.91% Mismatches: 264
Query Match: 4.84% Indels: 253
DB: 20 Gaps: 35

US-09-914-168-2 (1-919) x AAX91597 (1-2409)
Qy 300 ProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeuLe 319
Db 409 CCGTGTTCATCATCAGTCAGGTC-----AAATCCGATAGCGGCTAAGGTGGCG 456
Qy 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAsp 339
Db 457 ACGAATCTCCCGGAACACGCGGTACTTCGATGCTAA---GTAAAGACGATGTGACC 513
Qy 340 ValIleLeuProAspAsn---ThrAlaAspValSerLeuIleTyrAspThrGlyThrGln 358
Db 514 ACTCTGAAAAAGGACTCGCTCAAGCCCAAAATCTCTATACGGTGGATATGCGCTCTCT 573
Qy 359 TyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThr 378
Db 574 TATCATTCAGCACGATCATTCCTCCCTTACCGATC-----AGCACT 612
Qy 379 AspProAspLysLeuProValLysArgGluLeuLeuGluInLeuThrValAsnMet 398
Db 613 TTCCCGACACGATTCCTGGCTTACAGGCGACTCGCTCT-----TTGATCAGGAAA 663
Qy 399 GlyGluAlaTyrAsnLeu-----GlnAlaValArgAlaLeuSer 411
Db 664 GGAGACCAGTTCAATTTGGCAAGCTGCACGAAGAGCGTCAGACCATCAGTCCCTGCTG 723
Qy 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431
Db 724 AGAGAC-----AATGGTTACTACTTCTCCGCCACAGGATATTATCTACGAAGCC 774
Qy 432 ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGlu 451
Db 775 GATACCTCTCTCGTAAGAGGTGCGCTATGCTCGGAGCCAAAGCTCTCGGAAGATACTCCA 834
Qy 452 ProAla-----GlnValAspGluSerThrLeu----- 460
Db 835 CCCCAAGCCATGCGCCCGTGGAGGATAGGAAACGACAGCAGCTCTCTCGGAATGAAC 894
Qy 461 ---GluProValIleGluThrValGluLeuThrAsp---GlyIleLeu----- 474
Db 895 GGAGAAAGCCCGACAGACTCGCTCGAGTGAGGATATGAAGTCTTTTACTATCGTAAA 954
Qy 475 MetAspIleSerPro-----IleGluPheSerAlaSerAsnLeuLe 488
```

Qy	770	rGlyValTyrSerPheClyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGl	790
Db	1914	CGTGATATATAGCTATGGC-----	1932
Qy	790	yGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuAr	810
Db	1933	-----AATATGGCAGTGGCACCTATATCGGACCA	1961
Qy	810	gPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu-----	827
Db	1962	GTTCATGTAGGCGGTGCCAATAGTATCAGAGCTTTCCACCGTCCGTAGCATCGGCCCGG	2021
Qy	828	-----SerProIleSerAspLysGlyTyrLeu-----ThrGlyGlyGlnValLe	842
Db	2022	ACGGTTCATCCGGATTCGCGACAATCAGTATTCCCTATTGTGATCAGGTGGCGGAATTCAA	2081
Qy	842	uAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPh	862
Db	2082	ACTCGAAGCCACGCTGAATATAGAGGCAAGCTTTTCGGGGATCTCCACGACGCGCTTTT	2141
Qy	862	eGlyAspIleGlyAsnAlaTyr-----	869
Db	2142	CCTCGATCGCGCAAGCTTTGGCTCTTCAGGAGGAGTATCTCCCGTCGCGGCGGTGCTCT	2201
Qy	870	-----AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGln	884
Db	2202	GTCCGAAGTGGATCGGTGAGCAATTTCTCTGAATAGCATCGCTCTCGGCACCGGTGTGCG	2261
Qy	884	yValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys--	903
Db	2262	CCTTCGCTAGCATCTGCATTTCTCGTGTTCTGTCGATGTGCGGTTCGCTCCACCT	2321
Qy	904	-----GluGluGly-----AsnProIleLeu	910
Db	2322	TCCTTACAATACGGGTAGAAAGTTTACTACAATATCCACGCTTTAAGGATGCATCGG	2381
Qy	910	sLeuHisPhePheIleGlyThrProPhe	919
Db	2382	TTTCCATTTGGTGTGGCTATCCCTTC	2409
RESULT	24		
AAA48507	ID	AAA48507 standard; DNA; 1830 BP.	
XX	XX	AAA48507;	
DT	DT	03-JAN-2001 (first entry)	
XX	XX	Neisseria meningitidis BASB040 putative coding sequence #1.	
DE	DE	BASB040; bacterial disease; respiratory tract infection; bacteraemia;	
KW	KW	meningitis; cancer; autoimmune disease; ds.	
XX	XX	Neisseria meningitidis.	
OS	OS		
XX	XX		
FH	FH	Key Location/Qualifiers	
FT	FT	CDS 1..1830	
FT	FT	/tag= a	
FT	FT	/product= "BASB040"	
XX	XX		
XX	XX	WO200034480-A1.	
XX	XX	15-JUN-2000.	
XX	XX	02-DEC-1999; 99WO-EP09560.	
XX	XX	07-DEC-1998; 98GB-0026886.	
XX	XX		
PA	PA	(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.	
XX	XX		
PI	PI	Ruelle J;	

DR WPI: 2000-423426/36.  
 P-PSDB; AAY99623.

XX Novel BASB040 polypeptides of *Neisseria meningitidis* useful for  
 PT diagnostic, prophylactic and therapeutic purposes against microbial  
 PT diseases comprise a specific amino acid sequence -

XX  
 PS Claim 14; Page 59; 98pp; English.

XX The present sequence is a putative version of the *Neisseria*  
 CC meningitidis strain ATCC13090 BASB040 coding sequence. The protein  
 CC produced from this gene is similar to the B15 outer membrane protein of  
 CC the bacterium. The gene, its protein, antibodies, antagonists and  
 CC agonists can be used to diagnose and treat bacterial diseases such as  
 CC those leading to upper respiratory tract infections, bacteraemia and  
 CC meningitis. In addition, they can be used in vaccines for use against  
 CC cancer and autoimmune diseases.

XX  
 SQ Sequence 1830 BP; 440 A; 603 C; 461 G; 326 T; 0 other;

# Alignment Scores:

Pred. No.: 9.4e-09 Length: 1830  
 Score: 224.50 Matches: 135  
 Percent Similarity: 34.33% Conservative: 72  
 Best Local Similarity: 22.39% Mismatches: 218  
 Query Match: 4.75% Indels: 178  
 DB: 21 Gaps: 19

US-09-914-168-2 (1-919) x AAA48507 (1-1830)

QY 462 ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGlu 481  
 DB 124 CCGGACACCGAATCAGTAAATTAATAAACCACCAATTCGCCGTCGACATCGACACGAGAT 183  
 QY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501  
 DB 184 AGTGAATCAAGATATGGTCGAAGAACACCTCGCGCTCATCGACGACGAGGAGAA 243  
 QY 502 LeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg 521  
 DB 244 GTATTGGACAAAGGACAGCGGCTTCTCGCCGGAAGACCGGACACGTTAAACCC 303  
 QY 522 SerIle-----LeuGlyArgIleSerAspAlaValSerAla----- 533  
 DB 304 ATGCTCCGACGCAAGAGCTATTTCAGCAGCAAAAGTCAGCGCTCGAGGAAAGACGAGCT 363  
 QY 534 -----ValAlaArgAlaIle 538  
 DB 364 TATACGGTACATCACACCGGGCCCGCACCACCAAAATCGCCACGTCGGTGTCCGCATC 423  
 QY 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558  
 DB 424 CTCGGCAGACATCCTTTTCAGACGGCAACCTCGCGGAATACTACCGCAACGCGCTGGAAC 483  
 QY 559 ArgLysThrProAla-----AspValTyrGlnSerLys----- 570  
 DB 484 TGGCAGCAGCGGTAGGACGTATTTCGATCAGCAGCAGTTGGGAAACAGCAAACTTCC 543  
 QY 570 ----- 570  
 DB 544 GTCTCGCGGGGTAAACGGGAAAGAGGTACCGGCTTCCCAAGCTCGGCAACACCCGGGCG 603  
 QY 571 -----ValProLeuTyrValPheValAlaSerAspLysPro 582  
 DB 604 GCGGTCAACCCGATACCGCCACCGCTGATTGAACGTCGTGCGGACGCGCGGCC 663  
 QY 583 ArgAspGlyGlnIleGlyLeuGlyTyrP6LysSerAspThrGlyThrArg----- 598  
 DB 664 -----ATCGCCTTCGGGACATTTGAATATCACCGGCACACACGCTTACCCGGAA 711  
 QY 599 -----LeuVal 600  
 DB 712 CAATCTCTCCGGCCTGGCGCGCTTCCAAACCGGGGACGCCCTACGACCTCGACCTGCTG 771

QY 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu--- 619  
 DB 772 CTGCACTTCCAAACAGGCATC---GAACAAACCGGCATATTTCGGCGCGTCCGTACAA 828  
 QY 620 -----ArgLeuSerGluAspLysGlyValLysLeuTyrAlaThrLysPro 635  
 DB 829 GCGGACTTCGACCGCTCTCCAAAGCGACGCGTCCCGCTCAAAGTCAGCGTAACCGAGTC 888  
 QY 636 LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPhe 655  
 DB 889 AAGCGCCACAGCTCGAAACCGGCATCGCGCTCGATTTCGGGAATAC----- 933  
 QY 656 GlyHisSerThrAsnGly-----PheAsp-LeuSerThrArgThrLeuGluHisG1 672  
 DB 934 -----GTTTGGCGCGCAAAATCGCCTACGACTATTACAACTCTTCAACAAAGCTAT 987  
 QY 672 uIleSerArgSerIleIleGlnAsnGlyTyrTrpAsnArgThrTyrSerLeuArgTyrAr 692  
 DB 988 ATCGGCTCGGTCG-----TCTGGGATATGGACAAATACG----- 1021  
 QY 692 gLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPh 712  
 DB 1022 -----AAACACCGCTTTCGCCGCGCATCGCCAGCGCGCAACTATCGGGGC 1068  
 QY 712 eValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrVa 732  
 DB 1069 AACTACTTGA-----CAAGCAACGTTTCTTCAACCGCTTCGACACCAACCAACC-- 1117  
 QY 732 lAlaAspAsnLeuValAsnPro-----Me 740  
 DB 1118 -TCGAAACACGCGCTTCTCGCGCGCATCTGTATGTGCGCGACCGCGCGGCGCATCGAT 1176  
 QY 740 tArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSer--GlyLeuValSerAsp 759  
 DB 1177 GCGAGCTGGGGCAGAGTTTCTCGCAGAGCGCGGAAATATCCCGGC-----TCGGAT 1230  
 QY 760 AlaAsnMetAlaIleAlaArgAlaGlyIle-----SerGlyValTyrSerPheGly 776  
 DB 1231 ATCGATTGGGCAACAGCGACGCGCATGCTGACCGCTCTTGGAAACGCGCAGCTGCTC 1290  
 QY 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIle----- 792  
 DB 1291 AACACGCTGTGCATCCGAAACAGCGCATTTACCTCGACGCGCAAAATCGTACGACTTGG 1350  
 QY 792 -----j----- 792  
 DB 1351 GGCACATTCTCTCTCCACCGCGTGTATCGCAGCTCTGCGCGTGCAGGTTATTTCTTC 1410  
 QY 793 -----GlnAlaGlyTyrIleTrp 798  
 DB 1411 ACGCCGCAAAACAACTCGGACGCTTCATCATACGCGGCAACAGCGGTTACACCGTT 1470  
 QY 799 SerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSer 818  
 DB 1471 GCCCGCAACATCCGACGCTCTCTTCAGGCTGTATTTCCGACGCGCGCGCTCTTCC 1530  
 QY 819 IleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGly 838  
 DB 1531 GTGCGGTTACGAATCGACAGCATCGGACTTCCCGCGCCGCAACGAGTCTCTCGCC 1590  
 QY 839 GlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArg 858  
 DB 1591 GAACGCGCGCTCTCTGGGCGCGCTGGAATACCAACTGCGGTTTACGCGCAGCTTTC 1650  
 QY 859 LeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLys 878  
 DB 1651 GCGCGGTGTTCACGATATGGGCGCGCGCGCCCAATTTCAACAGCTATGAAGCTGAA 1710  
 QY 879 IleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArgValAspVal 898  
 DB 1711 CACGCTTCGGGACTGGCGCTGGTTCAGCGCGCTTTCCTTCCTTCGACATC 1770





```
Db 454 -----CTTTCAGACGGCAACCTCGCCGAATACTACCGCAACGCGCTGGAAACATGG 504
Qy 478 ---SerProIle-----GluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeu 494
Db 505 CAGCAGCGGTAGGACAGTTCGATCAGGACAGTTCGGAAACAGCAAACTTCCTGC 564
Qy 495 ValAlaAlaLysAlaArgHisLeuTyrAspMetPro-----AspAspArgValLeuAla 512
Db 565 CTCGGCGGGTAACGCGCAAGCCTACCGCTTGCAAGCTCGGCAACACCGCGCGGCC 624
Qy 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
Db 625 GTCACCCCGATGATGACC-----ACCGCGATTGAAAC 657
Qy 533 AlaValAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
Db 658 GTCGTCGGGACAGCGCGGCCCA-----TCGCCTTCGGCGACTTCGAAATTACCGGC 711
Qy 552 uArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValPr 572
Db 712 ACGCAGCGTTACCCGGAACAAATCGTCTCCGGCT----- 745
Qy 572 oLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGl 592
Db 746 -----TCGCGCGCTTCCACCGGCGACGC----- 769
Qy 592 ySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGl 612
Db 770 ---CCTACGACCTCGAC-CTGCTGCTCGACTTCCACACAGCGCTC---GAACAAACGG 821
Qy 612 yTyrGlnAlaGlyAlaGluLeu-----ArgLeuSerGluAspLysLysGl 627
Db 822 GCATTATTCCGCGCGCTCGGTACAGCGGACTTTCACGCGCTCCCAAGCGCGCTGCC 881
Qy 627 yValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaTh 647
Db 882 CGTCAAGTCACGTAACAGGTCAACAGCCCAAGCTCGAAACCGGCACTCCGCTCGCA 941
Qy 647 rLeuGlyTyrGlnGluValPheGlyHisSerThrAsnGly----- 661
Db 942 TTCGGAATAC-----GGTTGGCGCGCAAAATCGCTACGACTA 980
Qy 662 -PheAspLeu-SerThrArgThrLeuGluHisGluIleSerArgSerIleLeuGlnAsnG 681
Db 981 TTACAACCTCTTCAACAAAGGCTATA-----TCGGTTGCGTCTGCTG 1022
Qy 681 yGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaP 701
Db 1023 GGATATGGACAAATACG-----MetArgGlyTyrArgGlnArgTyrSerL 749
Qy 701 roProGluThrTrpGlnAspLeuProValAspPheValasnGlyLysProSerGlnGluA 721
Db 1053 CGCGGCACTACGCCAGCGCGCAACTATCGGGGCACTACTGGA-----CAAGCAA 1103
Qy 721 laLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnPro---- 739
Db 1104 CGTTTCTTACACCGCTTCGACCAACCAAAAC-----TCGAAAAACGCGCTTCTCGGGG 1160
Qy 740 -----MetArgGlyTyrArgGlnArgTyrSerL 749
Db 1161 CATCTGGTATGTCGCGACCGCGGGGATCGATCCAGGCTGGGGCGGAGTTTCTCGC 1220
Qy 749 euGluValGlySerSer--GlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGl 768
Db 1221 AGAAGCGCGGAAATCCCGCGC-----TCGGATATCGATTGGGCAACAGCCGCCAC 1274
Qy 768 yIle-----SerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAl 785
Db 1275 GATGCTGACCGCTCTTGGAAACCGCAGCTGCTCAACACGTCGTCATCCGCAAAACGG 1334
Qy 785 aHisGlnMetThrGlyGlyIle----- 792
Db 1335 CCATTACCTCGACGCAAAATCGGTACGACTTTGGCGCAATTCCTGTCTCCACGCGCT 1394
```

```
Qy 792 ----- 792
Db 1395 GATCCGCACCTCTGCCGCGCAGGTTATTTCTTACGCCCGGAAACAAAACTCGGCAC 1454
Qy 793 -----GlnAlaGlyTyr---IleTrpSerAspAsnPheAsnHisValPr 806
Db 1455 GTTCATCATACGCGACAAAGCGGTTACACCGTTTGCCTCCGCAATGCCAAC---GTTCC 1511
Qy 806 oTyrArgLeuArgPhePheAlaGlyClyAspGlnSerIleArgGlyTyrAlaHisaspSe 826
Db 1512 TTCAGGGCTGATGTTCCGACGCGCGCGCTTCCGTGCGCGGTTCAGAACTCGACAG 1571
Qy 826 rLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyTh 846
Db 1572 CATCGGGCTTCCGCGCGCAACGATCGTCTCTCCCGAACCGCCCTCTTGTGGGCGAG 1631
Qy 846 rAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGl 866
Db 1632 COTGAATACCAACTGCGCTTTTACGCGCACCTTTCCGCGCGGTGTTCCACGATATGG 1691
Qy 866 yAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValAr 886
Db 1692 CGACCGCGCGCAATTTCAACAGTATGAAGCTGAACACGCTCGGACTCGGCGTGGC 1751
Qy 886 gTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLysGluGluGl 906
Db 1752 CTGGTTACGCGCGCTCGCGCGCTTTTCTTCCACATCGCTACGCGCACGCGCAAG-- 1809
Qy 906 yAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 1810 ----AAAATCCGCTGGCATCATCAGCTTGGAAACGCGCTTC 1845
RESULT 27
AAA48509
ID AAA48509 standard; DNA; 1764 BP.
AC AAA48509;
XX
XX 03-JAN-2001 (first entry)
XX
XX Neisseria meningitidis BASB040 putative coding sequence.
XX
XX BASB040; bacterial disease; respiratory tract infection; bacteraemia;
XX meningitis; cancer; autoimmune disease; ds.
XX
XX Neisseria meningitidis.
XX
XX Key Location/Qualifiers
XX CDS 1..1764
XX FT /*tag= a
XX FT /product= "BASB040"
XX FT /partial
XX
XX WO200034480-A1.
XX
XX 15-JUN-2000.
XX
XX 02-DEC-1999; 99WO-EP09560.
XX
XX 07-DEC-1998; 98GB-0026886.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI; 2000-423426/36.
XX DR P-PSDB; AAY99625.
XX
XX Novel BASB040 polypeptides of Neisseria meningitidis useful for
XX diagnostic, prophylactic and therapeutic purposes against microbial
XX diseases comprise a specific amino acid sequence
XX
```

PS Claim 11; Page 61; 98pp; English.

CC The present sequence is a putative version of the *Neisseria meningitidis* strain H44/76 BASB040 coding sequence. The protein CC produced from this gene is similar to the D15 outer membrane protein of CC the bacterium. The gene, its protein, antibodies, antagonists and CC agonists can be used to diagnose and treat bacterial diseases such as CC those leading to upper respiratory tract infections, bacteraemia and CC meningitis. In addition, they can be used in vaccines for use against CC cancer and autoimmune diseases.

XX Sequence 1764 BP; 424 A; 581 C; 451 G; 308 T; 0 other;

#### Alignment Scores:

Pred. No.:	4.02e-08	Length:	1764
Score:	216.50	Matches:	136
Percent Similarity:	34.87%	Conservative:	76
Best Local Similarity:	22.37%	Mismatches:	208
Query Match:	4.58%	Indels:	188
DB:	21	Gaps:	21

US-09-914-168-2 (1-919) x AAA48509 (1-1764)

Qy	462	ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGlu	481
Db	58	CCCGACACCGAATCAGTCMAATTAACCCCAAAATCCCGCTCCTCATCGACACGAGAC	117
Qy	482	PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnValAlaAlaLysAlaArgHis	501
Db	118	AGTGAATCAAGATATGTCGAAGAACACCTCGCGTCTCATCGACGACGAGGAAGAA	177
Qy	502	LeuTyrAspMetProAspAspArgValIleLeuAlaIleAsnHisAspAspGlyValAsnArg	521
Db	178	GTATTGCAACAGGACAGCGGCTTCCTCGCCGAAGAGCGCGGACACAGTTTAAACG	237
Qy	522	SerIle-----LeuGlyArgIleSerAspAlaValSerAla-----	533
Db	238	ATGCTCGCAGCAAGAGCTATTTCAGCAGCAAAAGTCAGCTGACGGAAAAAGACGGAGCT	297
Qy	534	-----ValAlaArgAlaIle	538
Db	298	TATACGTTACATCATCACCGGCGCCGCCAACCAAAATCGCCAGCTCGCGCTGCCATC	357
Qy	539	LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn	558
Db	358	CTCGGCACATCTTTCAGACGGCAACTCGCCGAATACTACCGCAACGCGCTGGAAAC	417
Qy	559	ArgLysThrProAla-----AspValTyrGlnSerLys-----	570
Db	418	TGGCAGCAGCGGCTAGCAGCGATTTCGATCAGACAGATTGGGAACACAGCAAACTCC	477
Qy	570	-----	570
Db	478	GTCTCGCGCGGTAGCAGCGCAAGCCTACCGCTTGCCAAAGCTCGGCAATACGACGGG	537
Qy	571	-----ValProLeuTyrValPheValAlaSerAspLysPro	582
Db	538	GCGCTCAACCCCGATACCGCCACCGCGGATTTGAACGTCGCTGGACAGCGCGCGCC	597
Qy	583	ArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrArg-----	598
Db	598	-----ATCGCTTCGCGGCTTTGAATCACCAGGACACAGCGCTTACCCCGAA	645
Qy	599	-----LeuVal	600
Db	646	CAAAATGCTCTCCGCGCTTGCGGCTTTCCAGCCGCGGTATGCCGTACGACCTCGACCTGCTG	705
Qy	601	ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu---	619
Db	706	CTCGACTTCAACAGCGGCTC-----GAACAAAACGGGCTATTATTCGCGCGGCTCCGTACAA	762
Qy	620	-----ArgLeuSerGluAspLysGlyValLysLeuTyrAlaThrLysPro	635

Db	763	CCGACTTCGACCGCTCCAAAGGCGACCGCGTCCCGCTCAAGTCCAGGCTAACCCAGGTC	822
Qy	636	LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPhe	655
Db	823	AAAGCCACAAATCGAAACCGGCATCGCTCGATTCCGAATAC-----	867
Qy	656	GlyHisSerThrAsnGly-----PheAspLeu-SerThrArgThrLe	669
Db	868	-----GGTTTGGCGCGCAAAATCGCTACGACTATTACAACCTCTTCAACAAGGCTAT	921
Qy	669	uGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTyrSerLe	689
Db	922	A-----TCGGTTTCGCTCGGTATATGGACAATAACG-----	955
Qy	689	uArgTyrArgLeuAspLysLeuLys-----	697
Db	956	AAACACGCTTCGCGCGGCATCAGCCAGCGCGCAACTATCGGGGCACTACTGGACA	1014
Qy	698	-----ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPh	712
Db	1015	AGCAACGTTTCTTACACCGTTTCGACCAACCAACCTCGAAAAACGCGCTTCTCCGCG	1074
Qy	712	evalAsnGly-----LysProSe	718
Db	1075	GGCTCTGGTATGTCGGGACCGCGGCATCGATCCAGCTGGGGCGGGAATTTCTC	1134
Qy	718	rGlnGluAla--LeuLeuAlaGlyValAlaVal-----HisLysThr---	731
Db	1135	GCAGAAGCGCGGAAATCCCGCGCTCGGCTCGATTGGGCAACAGCAGCCGACGATG	1194
Qy	732	-----ValAlaAspAsnLeuValAsnProMetArgGlyTyr	743
Db	1195	CTGACCGCTTTCGAAACGCGCAGCTGCTCAACACGTCGTCATCCGCAAAACGCGCAT	1254
Qy	744	-----ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal	757
Db	1255	TACCTCGACGCCAAATATCGGTACGACTTTGGGCACATTCCTGTCTCCACCGCTGATC	1314
Qy	758	SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp	777
Db	1315	CGCACCTCT-----GCCCGTCGAGT-----	1335
Qy	778	AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIle-----Gln	793
Db	1336	-----TATTTCTTCACGCGCGCAAAACAAACAACTCGGACGTTTCATCATACGCGGACA	1389
Qy	794	AlaGlyTyrIleTrpSerAspAsnHisValProTyrArgLeuArgPhePheAla	813
Db	1390	GCGGTTTACCGTTGCGCGGACAAATCCGACAGCTTCTTCAGGGCTGATGTTCCGCGAGC	1449
Qy	814	GlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys	833
Db	1450	GCGCGCGCTTTCGCTGCGCGGTACGAACCTCGACAGCATCGGACCTTGGCGGCGCGAAC	1509
Qy	834	GlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPhe	853
Db	1510	GGATCGCTTCGCGCAACGCGCGCTCTCTGTGGGACGCTTGAATACCAACTGCCGTTT	1569
Qy	854	MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe	873
Db	1570	ACGGCACCTTTCCGCGCGGTGTTCCACGATATGGCGATGCCGCGCGCAATTTCCAA	1629
Qy	874	ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGln	893
Db	1630	CGTATGAAGCTGAACACGCTTCGGGACTCGGGCTGGCTGGTTCACCGCGCTCCGCGC	1689
Qy	894	ValArgValAspValAlaThrGlyValLysGluGlyAsnProIleLysLeuHisPhe	913
Db	1690	TTTTCCTTCACATCGCTACGGCAGCAGGATAAG-----AAAATCCGCTGGCACATC	1743
Qy	914	PheIleGlyThrProPhe	919
Db	1744	AGCTTGGGAACACGCTTC	1761







```
Db 38631 TATACGGTACATACACACCGGGCCCGCCACCAAAATCGCCAACTCGCGCTGCCCATC 38690
Qy 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558
Db 38691 CTCGGCAGCATCTTTTCAGACGGCAACTCGCCGAATACTACCGCAGCGCTGGAAAC 38750
Qy 559 ArgLysThrProAla-----AspValTyrGlnSerLysLys----- 570
Db 38751 TGGCAGCAGCGCGTAGGAGCGGATTCGATCAGGACAGTTGGGAAACAGCAAACTTCC 38810
Qy 570 ----- 570
Db 38811 GTCTCTGGCGGGTAACCGCGCAAGCCTTACCGCTTGCCAAAGTCGGCAATACGACGGC 38870
Qy 571 -----ValProLeuTyrValPheValAlaIleAspLysPro 582
Db 38871 GCCGTCAACCCCGATACCGCCAGCGGATTTGAACGTCGTCGTGGACAGCGCGGCC 38930
Qy 583 ArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArg----- 598
Db 38931 -----ATCGCTTCGGCGACTTTGAAATCACCGGCACACAGCGTTACCCCGAA 38978
Qy 599 -----LeuVal 600
Db 38979 CAATATGCTCCGCGCTTCGCGCTTTCAGCCCGGTATCGCGTACGACCTCGACCTGCTG 39038
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu--- 619
Db 39039 CTCGACTTCCAAACGGCGCTC--GAACAACACGGCGATTATTCGGCGCGCTCGGTACAA 39095
Qy 620 -----ArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysPro 635
Db 39096 GCCGACTTCGACCGCTCAAGGCGACCGCTCGCGCTCAAGTCAGCGTAACCGAGTCC 39155
Qy 636 LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPhe 655
Db 39156 AAACGCCCAAACTCGAAACCGGCATCCGCTCGCTCGGAAATC----- 39200
Qy 656 GlyHisSerThrAsnGly-----PheAspLeu-SerThrArgThrLe 669
Db 39201 -----GGTTTGGCGGCAAAATCGCTTACGACTATTACAACCTTCAACAAAGGCTAT 39254
Qy 669 uGluHisGluIleSerArgSerIleIleGlnAsnGlyClyTyrPsnArgThrTyrSerLe 689
Db 39255 A-----TCGGTTCGGTCTGCTGGGATATCGACAAATACG----- 39288
Qy 689 uArgTyrArgLeuAspLysLeuLys----- 697
Db 39289 -AAACACGCTTGGCGGGCATCAGCCAGCGCACTATCGGGCACTACTGCGACA 39347
Qy 698 -----ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPh 712
Db 39348 AGCAACGTTTCTTACACCGTTTCGACCAACCAACCTCGAAACACGCGCTTCTCCGCG 39407
Qy 712 eValAsnGly-----LysProSe 718
Db 39408 GCGCTCTGGTATGTCCGCGACCGCGCGGCATCGATGCCAGGCTGGGGCGGGAATTTCTC 39467
Qy 718 rGlnGluAla--LeuLeuAlaGlyValAlaVal-----HisLysThr--- 731
Db 39468 GCAGAGGCCGGAATATCCCGGCTCGCTGTGCTGTTGGCAACAGCAGCCACGCGATG 39527
Qy 732 -----ValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
Db 39528 CTGACCGCTCTTGGAAACGCCAGCTGTCTCAACAACGTGTCTCATCCCGGAAACGGCCAT 39587
Qy 744 -----ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal 757
Db 39588 TACCTCGACGCAAAATCGGTACGACTTTGGGCACATTTCTGCTCTCCACCGCGCTGATC 39647
Qy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 39648 CGCACCTCT-----GCCGTCAGGT----- 39668
```

---

```
Qy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIle-----Gln 793
Db 39669 -----TATTTCTTCCAGCCCGCAAAACAAAACAACTCGCGACGTTTCATCATCGCGCAAA 39722
Qy 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
Db 39723 CGCGGTTTACACCGTTTCGCGCGGCAAAATCCGACGCTTCTTCAGGCTGATGTTCGCGACG 39782
Qy 814 GlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833
Db 39783 GCGCGCGGCTTTCGCTCGCGGTTTACCACTCGACACATCGGACTTCCGCGCGCGCAAC 39842
Qy 834 GlyTyrThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPhe 853
Db 39843 GGATCGGTCCTCGCCGAACGCGCTCTCTGTGGGACGCTGGAATACCACTCGCGCTTT 39902
Qy 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe 873
Db 39903 ACGCGACCTTTCGCGCGGCTGTTCACGATATGGCGGATGGCGCGCCCAATTTCAAA 39962
Qy 874 ThrAsnAspThrLysIleGlyValGlyValArgTyrAlaSerProValGlyGln 893
Db 39963 CGTATGAAGCTGAACACGCTTCGGGACTGGCGTGGCTGTTTCAGCCCGCTTCGCGCG 40022
Qy 894 ValArgValAspValAlaThrGlyValLysGluGlyAsnProIleLysLeuHisPhe 913
Db 40023 TTTTCTTCGACATCGCTACGCGCACGCGATAAG-----AAATCCGCTGGCACATC 40076
Qy 914 PheIleGlyThrProPhe 919
Db 40077 AGCTTGGGAACGCGCTTC 40094
RESULT 30
AAF21613
ID AAF21613 standard; DNA; 172325 BP.
XX
AC AAF21613;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
diagnosis; antigen; detection; infection; gene therapy; antibacterial;
ds.
XX
OS Neisseria meningitidis.
XX
FN WO200066791-Al.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Piazza M, Hickey E, Peterson J, Tettelin H, Venter JC, Maignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI: 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7; Appendix A; 692pp; English.
XX
```



QY 874 ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGln 893  
 Db 143630 CGTATGAAGCTGAACACGGTTTCGGGACTGGCGTGGCTGTTCAGCCGGTTTCGGCGG 143689  
 QY 894 ValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisPhe 913  
 Db 143690 TTTCTCTTCGACATCGCTCCGACGACGCGATAAG-----AAAATCGCTGGCACATC 143743  
 QY 914 PheIleGlyThrProPhe 919  
 Db 143744 AGCTTGGGAACGGCTTC 143761  
 RESULT 31  
 ABO72980  
 ID ABO72980 standard; DNA; 2388 BP.  
 XX ABO72980;  
 AC ABO72980;  
 XX 20-SEP-2002 (first entry)  
 DE Methylococcus capsulatus surface protein-antigen D15 DNA SEQ ID NO:5.  
 KW Methylococcus capsulatus; methanotropic bacterium; MopC; MopD; MopE;  
 MW MopF; D15; surface protein antigen; helper protein; vaccine;  
 KW epitope mapping; gene; ds.  
 XX  
 OS Methylococcus capsulatus.  
 PN WO200255549-A2.  
 PD 18-JUL-2002.  
 PF 14-JAN-2002; 2002WO-NO00018.  
 PR 12-JAN-2001; 2001NO-0000238.  
 XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.  
 PA (TIGR-) TIGR.  
 XX Birkeland NK, Eldhammer I, Jonassen I, Jensen HB, Lien T;  
 PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;  
 PI Salzberg SL;  
 XX WPI; 2002-557813/59.  
 XX New nucleotide molecule encoding a surface exposed protein and is  
 PT incorporated and expressed in Methylococcus capsulatus, useful in  
 PT manufacturing vaccines that can be administered orally in animals, fish  
 PT or humans  
 PS Claim 1; Page 34-35; 39pp; English.  
 CC The present invention describes a nucleotide molecule (I) comprising a  
 CC sequence that codes for any of the 14 Methylococcus capsulatus  
 CC nucleotide sequences given in ABO72976 to ABO72989. Also described is  
 CC a method for producing a desired protein in a bacterial host cell,  
 CC comprising: (a) transforming a bacterial host cell with a recombinant  
 CC vector comprising a first nucleotide sequence selected from the  
 CC group of 14 Methylococcus capsulatus, and comprising a nucleotide  
 CC encoding the desired protein linked in frame to the first nucleotide  
 CC sequence; and (b) culturing the transformed host cell in a medium under  
 CC conditions that allow the expression of the protein. (I) can be used  
 CC in vaccine production. The nucleic acids or the proteins from the  
 CC present invention can be used in manufacturing vaccines that can be  
 CC administered in animals, fish or humans. The method is useful in  
 CC producing fusion proteins carrying foreign peptide sequences that can  
 CC also be useful in epitope mapping, screening of antibody libraries and  
 CC as immunogens. The present sequence represents the specifically claimed  
 CC Methylococcus capsulatus surface protein-antigen D15 nucleotide sequence,  
 CC from the present invention.  
 XX Sequence 2388 BP; 535 A; 653 C; 686 G; 514 T; 0 other;

Alignment Scores: 8.7e-08 Length: 2388  
 Pred. No.: 214.50 Matches: 187  
 Score: 32.31% Conservative: 108  
 Percent Similarity: 20.48% Mismatches: 316  
 Best Local Similarity: 4.54% Indels: 303  
 Query Match: 24 Gaps: 43  
 DB:  
 US-09-914-168-2 (1-919) x ABO72980 (1-2388)  
 QY 165 IleLysArgLeuThrAlaArgLeuPheAsnAsp-----GlyVal 177  
 Db 34 ATCCCATCGCTTCGGCGCAGGCCCTTGAACGATAGTTCTGCCGCATTTTTCAGGGGAGTG 93  
 QY 178 AsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSer 197  
 Db 94 GAATCCGTCGCGCCGCTGTGTCATGCTGCTGCTCATCCGGGACATCCGGATCGAG 153  
 QY 198 AlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGlu 217  
 Db 154 GCGTTC-----GAGCCCTTCGTCTCATCCGGGACATCCGGATCGAG 192  
 QY 218 AspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThr 237  
 Db 193 GGGCTGCACGGGATTTTCGGAGGGTACGGTGTTCACACTACCTGCCTGCTCGGGAGCGGAC 252  
 QY 238 AlaLeuValAlaAlaArgAla-----ValGlyTyrTyr 248  
 Db 253 ACCCTCGATGAAAACGGTGGCGGAGGTGATCCGGGCGCTGTTCAGACGGGTTTTC 312  
 QY 249 ---AspIleAspLeuSer-----IleIle-----ArgAsn 257  
 Db 313 AAGGACGTCGCTCGACGAGGACGAGCGTAAGTTGATCATCTACGTCTGAAGAGCGTCT 372  
 QY 258 SerIleGlyGluValAspValIle---IleHisAspLeuGlyGluProValTyrIleAsp 276  
 Db 373 TCGATTTCACGCTAAAGATCGACGCAACCATCATCGTACGAGGAGGATCTGCTGAAG 432  
 QY 277 TyrArgAlaValGluValArgGlyGlyGly---AlaAspAspLysAlaPhe---ThrThr 294  
 Db 433 -----CGGCTCAAGGGAATCGCGCTGGCGGAGGAGGATATTCAGCCGGCAG 480  
 QY 295 ValAlaAspGluValProLeu---LeuIleGlyAspValPheHisHisGlyLysTyrGlu 313  
 Db 481 ATTCTCGACAAGTTCGACAGGAACCTGCGCGCCAGTATTACAGCCGGGCAAAATACAGC 540  
 QY 314 ThrLysLysAsnIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrp 333  
 Db 541 CTAA--ATCGATTCCAGGTGACGGAA----- 567  
 QY 334 LeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyr 353  
 Db 568 -----CTGCCCCGGAACCGTGTGGCGGTGAATCAATATC 603  
 QY 354 AspThrGlyThrGlnTyrArgPheAspGluValAlaPhe-----PheThr 368  
 Db 604 GCGGAGGGCGCTTCCTCGCATAAACAGATCAATATCATCGGAATAATGCTATTAGT 663  
 QY 369 IleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGlu 388  
 Db 664 GATGATGCTCGACCCAGGATTT----- 687  
 QY 389 LeuLeuGluLeuLeuLeuThrValAsnMetGlyGluAlaTyr-----AsnLeuGln 405  
 Db 688 -----GAGCTGAGTACATCCATCTCTTCTATACCAAGGATGATCAT 738  
 QY 406 AlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThr 425  
 Db 739 TCCAGCAGAAATTTTCGGCCGATTCGGAGCGGCTCGCTCTCTACTATCTCTCGATCCGGGT 798  
 QY 426 GluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer 445  
 Db 445 ----- 445

```
Db 799 TACGTCAATTC-----GAAATCGAGTCCACCCAGGTATCGATCAGCCCAAC 846
Qy 446 SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465
Db 847 AAG-----AAAGAAATATATATC 864
Qy 466 ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer 485
Db 865 ACCATCAACGTTAAGGAGGCGCAGTCTTCAAGGTGGACGAGGTGACGGTCAACGCAAG 924
Qy 486 AsnLeuIle-----GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyr 503
Db 925 ACCATCGTCCACCCGACAACTCGTCCGCTCGTGGTATCGGCGCGGAGATATTTT 984
Qy 504 AspMetProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
Db 985 -----TCCAGGAAATGGCGACAGAAACCCAGAGGCG----- 1017
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
Db 1018 -----ATTCCGATCGTTGGCG----- 1095
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 1036 -----GAGGAGGCTATATCTTCGCCCAACGTCAATATGTCGCG 1074
Qy 564 AspValTyrGlnSerLysLysVal-----Pro 572
Db 1075 GACATCAATCAGGAGAAGACCGGTAAACATCACCTTTTCGTCGATCCGGCAACAG 1134
Qy 573 LeuTyrVal-----PheValAlaSerAspLysProArgAsp----- 584
Db 1135 GTCTACGTGAGGCGAATCAATTTCCAGGGCAATACCAGACCGGTGACGAGTGTGCGCG 1194
Qy 585 -----GlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 1195 AGGCAATTCGGCAGATGGAGCGCGCTGGCGTCG-----ACCGCG 1236
Qy 602 LysPheGluHisAsn-----LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu 619
Db 1237 AAGATCGAGCGTTCCAAAGACCAGCGTCTCGAGGCTCTCGGCTATTTCCAGGATGTCAACGTG 1296
Qy 620 ArgLeu-----SerGluAspLysLysGlyValLysLeuTyrAlaThrLys 634
Db 1297 GAGAGCGCGCGTGGCGGAGAACGAGTACGATCGATCGACGTCAACTACAGTGTGACCGAG 1356
Qy 635 ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr----- 650
Db 1357 AAGTCTTCC-----GGCAACTGACGCGCGGTGTAGGTATTTCCCACTCCAGTCCAG 1404
Qy 651 -----GlnGlnGluValPheGlyHisSer----- 658
Db 1405 GGCATCATATTCAATGCGCGGTCAACCCAGGACAACATTTTTCGGCAGCGGCAAGCGTGC 1464
Qy 658 ----- 658
Db 1465 AGTTCAATTTCAACACAGCCAGCATCAACACCATCTATGCTTGGGTATTTCAATCCT 1524
Qy 659 -----ThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671
Db 1525 TACGCCACCCCTGGATGCGATCGACAGTGGTTTCGACATCAGCTATCGCGACACCAACACC 1584
Qy 672 GluIleSerArg-----SerIleIleGlnAsnGlyGly---Trp 683
Db 1585 GGCATTTCGAACTATGTTGCCAACTACATTAACCAATGTGTTCCAGGTGGCGGCAACTGG 1644
Qy 684 AsnArgThrTyr-----SerLeuArgTyrArgLeuAsp-----LysLeuLys 697
Db 1645 GGTGTGCCATCGGAGAGTTCGACAGTATTCGGACCAACCTGGACTACAGTAACACCAAG 1704
Qy 698 -ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPr 717
Db 1705 CTCAAGACGACCTCC-----AGTCTTCGATCAATCAGACGCTTCATTGCC 1752
```

```
Qy 717 oSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVa 737
Db 1753 GACCATGGCAGCAGGATATTCGACCTATTCCTCCAC-----TGGGT 1794
Qy 737 lAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVa 757
Db 1795 TGGACCCCATACACCTGAATCGGCGCATATTCG----- 1828
Qy 757 lSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe----- 775
Db 1829 -----CGACGAGCGGTGGCGCGCAGCA-TTGACGGGACTGTTTGGCTCCCGTT 1877
Qy 775 ----- 775
Db 1878 CAGCACCTCGAGTATTACAAGGCCAACGTGCGCTGCGAGCAGTATTTTCGCTGACCCA 1937
Qy 776 -----GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMe 788
Db 1938 GGATCTGACGCTGTTGAATGGCATTTTCGGTTTACGG----- 1977
Qy 788 tThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrAr 808
Db 1978 ---GGCGCTATGCGAGTGC-----GGCAACACGCTCCTTCGTTCTCG 2018
Qy 808 gLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSe 828
Db 2019 GGAACACTTTTACGCGCGGTGCGAATTCGGTGGGTGATACCAAGCCCAATTCACCTGG 2078
Qy 828 rProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGl 848
Db 2079 GCCAGGACAGCCGTGGCTACGGCTTCGCGCGGCAACAGCAACTGACCGGCTCGGTGGA 2138
Qy 848 uTyrAsnTyr-----GluPheMetLysAspLeuArgLeuAlaValPh 862
Db 2139 GTTGTGTTTCCGTCGCTTTCGGGCGAGAACTGAAGAGTGTCCGCTTGGGTACCTT 2198
Qy 862 eGlyAspIleGlyAsnAlaTyr---AspLysGlyPheThrAsnAspThrLysIleGlyAl 881
Db 2199 CGTCGACGCGCGGAGTGTTCGTCATTCGCCCCAGCTTTCGGACCTTCGCTTCGAC 2258
Qy 881 aGlyValGlyValArgTyrAlaSerProValGlyGlnValArgValAlaThrGl 901
Db 2259 CGGATTTCCGCCAAGTGGCTATCACCTTCGGGCGCTGATGTCTCGATCGCCCGCC 2318
Qy 901 yValLysGluGlyAsnProIleLysLeuHisPhe 913
Db 2319 GCTCAACTCCAGAGCGGGGAGCGGAATTCAGCAATTTT 2355
RESULT 32
ABQ90024
ID ABQ90024 standard; DNA; 2391 BP.
XX
AC ABQ90024;
XX
DT 01-OCT-2002 (first entry)
XX
DE M. capsulatus gene #9 for DNA array.
XX
KW Micro array; gene; ds; differential expression; gene expression.
XX
OS Methylococcus capsulatus.
XX
PN WO200255655-A2.
XX
PD 18-JUL-2002.
XX
PF 14-JAN-2002; 2002WO-NO00019.
XX
PR 12-JAN-2001; 2001NO-0000235.
XX
PR 12-JAN-2001; 2001NO-0000239.
XX
PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
```



Qy	651	----	-----GlnGlnGluValPheGlyHisSer-	658
Db	1405	GGCATCATATTCAATCGCGCGTCA	CCCGAGCAACAATTTTCGGCAGCGGCAAGCGTGC	1464
Qy	658	----	-----	658
Db	1465	AGTTTCAATTTCACACACACCGAGATCA	CAACACCATCTATGCTTGGGGTATTTCACATCCT	1524
Qy	659	----	-----ThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis	671
Db	1525	TACGCCACCTGGATGGCATCAGCAGTGGT	TTTCGACATCAGCTATCGGACACCAACACC	1584
Qy	672	GluIleSerArg-	-----SerIleLeGlnAsnGlyGly----Tirp	683
Db	1585	GGCTATTTCGAACATATGTGCCAACTACA	TTCACCAATGTGTTCCAGTGGCGCGCACTGG	1644
Qy	684	AsnArgThrTyr-	-----SerLeuArgTyrArgLeuAsp-	697
Db	1645	GGGTGCGCCATCGGAGAGTTCGACAGAT	TATATATATATATATATATATATATATATATAT	1704
Qy	698	-ThrGlnAlaProGluThrTirpGlnAspLeu	ProValAspPheValAsnGlyLysPr	717
Db	1705	CTCAGAGCAGCTCCC	-----AGTCTTCGATCAAAATCAGAGCGTTCATTCGC	1752
Qy	717	oSerGlnGluAlaLeuLeuAlaGlyValAla	ValHisLysThrValAlaAspAsnLeuVa	737
Db	1753	GACCATGGCAGCGAGTATTCGACCTATTC	GCCTCTCAC-----TGGGT	1794
Qy	737	lAsnProMetArgGlyTyrArgGlnArgTyr	SerLeuGluValGlySerSerGlyLeuVa	757
Db	1795	TGGACCATGACACCCGTGAATCGGCGAT	TATTCG-----	1828
Qy	757	lSerAspAlaAsnMetalIeAlaArgAlaGly	IleSerGlyValTyrSerPhe-----	775
Db	1829	-----CGACCAGCGGTGGCGCAGCA-	-TTGACGGAGCTGTTGCCTGCCGTT	1877
Qy	775	----	-----	775
Db	1878	CAGCACCTCGAGTATTACAAGGCCAACGT	GCGCCCTGGAGCAGATTTTCGGCTGACCCA	1937
Qy	776	----	-----GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMe	788
Db	1938	GGATCTGACGCTGTGGTTGAATGGCGGAT	TTCGGTTACGGC-----	1977
Qy	788	tThrGlyGlyIleGlnAlaGlyTyrIleTirp	SerAspAsnPheAsnHisValProTyrAr	808
Db	1978	----GGCGGCTATGGCAGTGGC-	-----GGCAACAGCGTCTCTCGGTTCTG	2018
Qy	808	gLeuArgPheAlaGlyGlyAspGlnSerIle	ArgGlyTyrAlaHisAspSerLeuSe	828
Db	2019	GGAACACTTTTACCGCGCGGTCCGAATTC	CGTGGCTGGATACCGCCAAATTCACCTGGG	2078
Qy	828	rProIleSerAspLysGlyTyrLeuThrGly	GlyGlnValLeuAlaValGlyThrAlaGl	848
Db	2079	GCOCAGGACAGCGGTGGGTACGCCCTTC	GGCGGGCAACAGCAAACTGACCGCGTCGGTGA	2138
Qy	848	uTyrAsnTyr-	-----GluPheMetLysAspLeuArgLeuAlaValPh	862
Db	2139	GTGCTGTTCCGCTCCCGTTCGCGGGCAG	AAACTGAAGAGTGTCCGCTTGGGTACCTT	2198
Qy	862	eGlyAspIleGlyAsnAlaTyr---	AspLysGlyPheThrAsnAspThrLysIleGlyAl	881
Db	2199	CGTCGAGCGCGGAATGTGTTCGCAATTC	GCCCGCCAGCTTTCGGACCTGCGCTCTCGAC	2258
Qy	881	aglyValGlyValArgTirpAlaSerPro	ValGlyGlnValArgValAspValAlaThrGl	901
Db	2259	CGGCAATTCGCCCAAGTGGGTATACCCCT	TTCGGGCCCTGATGTCTCGATCGGCCAGCC	2318
Qy	901	yValLysGluGluGlyAsnProIleLysLeu	HisPhe	913
Db	2319	GCTCAACTCCAGACGGGGACCGGAATTC	AGCATTTT	2355



Db 91 TTGACTTCTCCAAAAGAGCCTCTCAAGAAATCTCAAAA-----AATGAAGCTCCA 141  
Qy 44 AlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrPro 63  
Db 142 AAAATGAAGTTCAAGAAATGAAGCTCAAAAAGAAACCCCAATCAATCAACGGCT 201  
Qy 64 GluGlnIleGlnAlaArg---LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGln 82  
Db 202 AAAGAAATGAAGTCAAGTCCATTTCTTATGTCGGGCTTTCTTACATCTCTGACATGCTC 261  
Qy 83 AlaLeuAspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSer 102  
Db 262 GCTAATGAATTTTAAAGATT-----CGTGGGCGCATATTGTG 300  
Qy 103 ProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGlu 122  
Db 301 GATTCTAAAAAATAGACACCGCTGTTTG-----GCT 333  
Qy 123 LeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyr 142  
Db 334 TTGTTCAATCAA----- 345  
Qy 143 GlnGlyGluGlnProAsnSerGluValValProProThrLeuGluProGluLysPro 162  
Db 345 ----- 345  
Qy 163 GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArg 182  
Db 346 GGGTATTTTAAAGACGTTTATGCC---ACTTTGAGCGCGCAT----- 387  
Qy 183 LeuLysAlaLysPheTyrGlnSerSerGlnSerGly-----GluThrSerAlaIleGly 200  
Db 388 TTAGAGTTTCATTTTGATGAAAAAGCCAGGATTGCGGGGTAGAAATCAAGGGTTATGGG 447  
Qy 201 SerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThr 220  
Db 448 ACTGAAAGGAAAAAGACGGCTTAAATCCCAATGGGGATCAAAAAGGCGCACACCTTT 507  
Qy 221 GlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu--- 239  
Db 508 GATGAGCAAAAATTAGAG-----CATGCTAAAACGGCITTAAAA 546  
Qy 240 ValAlaAlaArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIle 259  
Db 547 ACCGCTTTAGAGGGCGACGGCTATTAT----- 573  
Qy 260 GlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAla 279  
Db 574 GGGACCGTG----- 582  
Qy 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAsp----- 297  
Db 583 GTGGAGGTGCGACAGAAAAGTCAGTGAGGTGCATTATTGATCGTGTTCATGTGAAT 642  
Qy 298 ---GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys 316  
Db 643 AGGGGGGTAGCATTTATCAACAATCCATTTTATAGGGAAGCGCAAAATTAAGACGC 702  
Qy 317 AsnLeuIleGluAsnAlaSerAlaGluHis-----Gly 327  
Db 703 CGCATGTTGAATCTTTGAGTGCAGAACAGCAACAGAGATTTCATGGCTGGATGCGGGC 762  
Qy 328 TyrPheAspGlyArg----- 332  
Db 763 TTGAATGACGGAAATTCGGTTTAGATCACTAGAAATACGATTCTATCGGTATCCAAGAT 822  
Qy 333 -----TriLeuAspArgSerValasp-----ValIleLeuProAsp 344  
Db 823 GTGTATATCGGTAGGGTTTACTTAGACGCTCATATTCTTCGCCCTTTTGTAAAAACGGAT 882  
Qy 345 ---AsnThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGlnTyrArgPhe 361  
Db 883 TTTTCTACCCATGACGCTTAAGCTTCATTTATATAAGTCAAGAGGCGGATCCAATACAGGATT 942

Qy 362 AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAsp 381  
Db 943 TCAGACATTTTA-----ATAGACATTGACAAACCCG 972  
Qy 382 LysLeuProValLysArgGluLeuGluGlnLeuLeuThrValAsnMetGlyGluAla 401  
Db 973 GTAGTCCCTTAAAA-----ACCTTAGAAAAGCGCTTAAAGTGAAGGAAAGATGTC 1026  
Qy 402 TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu-----IleAlaThrArg 418  
Db 1027 TTTAATATGTAGCATTTAAGACGGATCGCAAAATTTTAAAAACCAATCGCCGATAAG 1086  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1087 GGTATGCGTTTGGC-----GTGCTGAACGACAGCTGGATAAAA----- 1125  
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSer 458  
Db 1126 -----CATGAAAAA 1134  
Qy 459 Thr---LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp--- 476  
Db 1135 AACGGGCTGTGAAGTTCATTTATCGTATTGAAGTGGCGGATATGGTGTATATCAATGAT 1194  
Qy 477 -----IleSerProIleGluPheSerAlaSerAsnLeuIle----- 488  
Db 1195 GTCATCATTTTCAGGAAACCGCACGAGCGATAGGATCATTAGAAGGAGTATTATTGTA 1254  
Qy 489 -----GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506  
Db 1255 GGGCTAAGGATAAATACAACCTTGACCAAACTGAGA----- 1290  
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526  
Db 1291 -----AATCCGAAAATCTTTAAGCGCT-----TTAGGATTC 1323  
Qy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546  
Db 1324 TTCCTAAAGTCAAAATTTGAAGAAAAAGGTT-----AATAGCTCA 1365  
Qy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566  
Db 1366 CTCATGGATTTA----- 1377  
Qy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586  
Db 1378 -----TTAGTGCCGTAGAAAGCGGCGTACTCGGCAG 1410  
Qy 587 Ile-----GlyLeuGlyTyrpGlySerAspThrGlyThrArgLeuValThrLysPheGlu 604  
Db 1411 TTGCAATTTGGTTAGGCTATGGCTCTTATGGAGGCTTATGCTT-----AAT 1458  
Qy 605 HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAsp 624  
Db 1459 GGGAGCTTGACGCAAGAAACCTTTTGGCACAGCGAAACGATAGCTTGTATGCTAAC 1518  
Qy 625 -----LysLysGlyValLysLeuTyrAlaThrLysPro----- 635  
Db 1519 ATCGCTACAGGGGGGTAGATCTTATCCGGGCATGCCAAAAGGCGGGCGTATGTTT 1578  
Qy 636 -----LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650  
Db 1579 GCCGGGAATTTGAGCTTGACTAATCCA----- 1605  
Qy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeu--- 669  
Db 1606 -----AGGATTTT-----GACAGCTGGTATAGCTCTACGATCAACCTTTAT 1647  
Qy 670 ---GluHisGluIleSerArgSerIleIleGlnAsnGlyGly----- 682  
Db 1648 GCGGATTACAGATAAAGCTACCAATACATACCAAGCGGGGGCTTTGGGGTGAATGTC 1707

```

QY 683 -----TrpAsnArgThrTyr---SerLeuArgTyrArgLeuAspLysLeuLys 697
Db 1708 GGCGCGATGCTGGGTAAAGAACCCATGTGAGCTTGGTAACTGAATGTACCAA 1767
QY 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValasnGly----- 715
Db 1768 CTCTTGGTTTACAGACCCCTTTATACACCGCTACTATTCTCTGTATGAAGTGGTT 1827
QY 716 LysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAla----- 733
Db 1828 TCTCCAAGCAATGTTCTACCCCGCATCGTGTATATCAATCGCTATCAGCGGTAAA 1887
QY 734 -----AspAsnLeuValAsnPro-----Met 740
Db 1888 ACCCCCTTACAACTGAAGCTGTTCTAGTCTGGAGCATCACCACTTCCACGAGAATA 1947
QY 741 ArgGlyTyrArgGlnArg-----TyrSerLeuGluVal 751
Db 1948 AGAGGTATTGGGATAGGGATTACCATACCCCTATCACAGCTCTTTCACCCCTTGATGTG 2007
QY 752 Gly-----SerSerGlyLeuValSerAspAlaAsn 761
Db 2008 AGCTATGACAAACCGATGATTATTACTTCCCTAGAAATGGGCTTATCTTTAGT----- 2061
QY 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp----- 777
Db 2062 -----TCCATCGCGAGATGTCTGGCTTCGCAAGCTCTGCACGCTCAATTTCTGG 2112
QY 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791
Db 2113 AACGGTTAGCGCGAATGTCCCTACACCAAAAGTTATGGTAAATCGCGCTTACCAC 2172
QY 792 -----IleGlnAlaGlyTyr 796
Db 2173 CATTTGCAAAATATTATTATGATAGATTGATCGCTGCTTTAAACGCAAGGAGGTAT 2232
QY 797 IleTrpSerAspAsnPhe---AsnHisValProTyrArgLeuArgPhePheAlaGlyGly 815
Db 2233 ATCTTAGGTATACACCGCATGATTACTTGCCTTAAACTCCACCTCTACATGGGGGGC 2292
QY 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyr 835
Db 2293 GTAACACCGTGAGGCTTTAGAACGGATCGTTACTCTCTAAAGATGATTTGCGTTG 2352
QY 836 LeuThrGlyGlnValLeuAlaValAlaGlyThrAlaGluTyrAsnTyrGluPheMetLys 855
Db 2353 TGGCTTGGAGCGCATGGGATTTTACCCTCTCTACTGAATTGAGCTATGGGTGCTAAAG 2412
QY 856 -----AspLeuArgLeuAlaValPheGlyAspIleGly----- 866
Db 2413 GCGGCTAAATGCGCTAGCGTGGTTTTTGTACTTTGGTTTCTTAACCTTTAAACCCCA 2472
QY 867 -----Asn 867
Db 2473 ACTAGAGGAGTTTTTCTATACGCTCTGTTACGACAGCGAATTTTAAAGATTATGGC 2532
QY 868 AlaTyrAspLysGlyPheThrAsnAspThr---LysIleGlyAlaGlyValGlyValArg 886
Db 2533 GTTATAGGGCTGGTGTGAAGACGCACTTGGAGGCTTCCACAGGCTTGCAGATTGAA 2592
QY 887 TrpAlaSerProValGly 892
Db 2593 TGGATTGCGCCCATGGGG 2610
RESULT 34
AAV52012
ID AAV52012 standard; DNA; 3012 BP.
XX
AC AAV52012;
XX
DT 09-NOV-1998 (first entry)
XX
DE Helicobacter polypeptide GHPO 107 DNA.

```

```

XX GHPO 107; infection; therapy; diagnosis; vaccine; gastritis; ulcer;
KW ss.
XX Helicobacter pylori.
XX OS
XX Key Location/Qualifiers
FH CDS 142..2685
FT /*tag= a
XX
XX W09821225-A1.
XX
XX 22-MAY-1998.
XX
XX 14-NOV-1997; 97WO-US21353.
XX
XX 29-JUL-1997; 97US-0902615.
XX 14-NOV-1996; 96US-0749051.
XX 01-APR-1997; 97US-0831309.
XX 01-APR-1997; 97US-0833457.
XX 01-APR-1997; 97US-0834705.
XX 24-JUN-1997; 97US-0881227.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
XX Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C;
XX Odenbreit S, Tomb J;
XX WPI; 1998-297855/26.
XX P-PSDB; AAW71477.
XX
XX Helicobacter polynucleotide and polypeptide sequences - useful to
XX treat or prevent gastrointestinal infection
XX
XX Disclosure; Page 82-86; 362pp; English.
XX
XX This polynucleotide encodes Helicobacter pylori polypeptide
XX GHPO 107 (see AAW71477). 85 Polynucleotides (see AAV52009-V52093)
XX encoding claimed Helicobacter polypeptides (see AAW71474-W71558) have
XX been identified in the H. pylori genome using a transposon shuttle
XX mutagenesis method in which TnMax9 mini-blam was used for
XX insertional mutagenesis of an H. pylori gene library established in
XX Escherichia coli. The invention provides: methods for producing
XX these polypeptides in recombinant host systems, and related
XX expression cassettes, vectors and transformed or transfected host
XX cells; live vaccine vectors that contain polynucleotides of the
XX invention and which can be used to prevent or treat Helicobacter
XX infection; therapeutic and/or prophylactic methods involving
XX administration of polynucleotide molecules (either in naked form
XX or delivered with a delivery vehicle), polypeptides or monospecific
XX antibodies; methods for detecting the presence of Helicobacter
XX in samples using e.g. polynucleotide molecules; and methods for
XX purifying polypeptides of the invention.
XX
XX Sequence 3012 BP; 896 A; 555 C; 735 G; 826 T; 0 other:

```

```

Alignment Scores:
Pred. No.: 6,39e-07 Length: 3012
Score: 205.50 Matches: 201
Percent Similarity: 32.89% Conservative: 143
Best Local Similarity: 19.22% Mismatches: 319
Query Match: 4.35% Indels: 383
DB: 19 Gaps: 51

```

US-09-914-168-2 (1-919) x AAV52012 (1-3012)

```

QY 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIleIleAsnHisValPro 43
XX ::::::::::: ||| ||| ::
XX
Db 25 TTGACTTCCCAAAAGAGCGCTCTCAAGATCTCAAAAA-----AATGAAGCTCCA 75
XX
QY 44 AlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrPro 63
XX

```

```
Db 76 AAAATGAAGTTCAACTCAAAAGAAAGAAACCCCAAAACCCCAATCAATCAAAACGCCT 135
    ::::: ::::: |||||
Qy 64 GluGlnIleGlnAlaArg--LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGln 82
    ::::: ::::: |||||
Db 136 AAAGAAATGAAGTCAAGTCCATCTTATGTCGGGCTTTCTTACATGCTGACATGCTC 195
    ::::: ::::: |||||
Qy 83 AlaLeuAspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSer 102
    ::::: ::::: |||||
Db 196 GCTAATGAATTTAAAGATT-----CGTGTGGCGCATATTCTG 234
    :::::
Qy 103 ProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 122
    ::::: ::::: |||||
Db 235 GATTCTAAAAAATAGACACCGCTGTTTG-----GCT 267
    :::::
Qy 123 LeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyr 142
    ||||| |||
Db 268 TTGTTCATCAA----- 279
    |||||
Qy 143 GlnGlyGluGlnProAsnSerGluValValProProThrLeuGluProGluLysPro 162
    |||||
Db 279 ----- 279
    |||||
Qy 163 GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAlaLysValProArg 182
    ||||| ::::: |||||
Db 280 GGGTATTTTAAAGACGTTTATGCC--ACTTTTGAAGCGGCATA----- 321
    |||||
Qy 183 LeuLysAlaLysPheTyrGlnSerSerGlnSerGly-----GluThrSerAlaIleGly 200
    ||||| ::::: |||||
Db 322 TTAGAGTTTCATTTTGATCAAAAGCCAGGATTCGGGGGTAGAAATCAAGGGTTATGGG 381
    |||||
Qy 201 SerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThr 220
    ::::: ::::: |||||
Db 382 ACTGAAAAGGAAAAGACGGCTTAAATCCCAATCGGGATCAAAAAGGGCCACACCTTT 441
    :::::
Qy 221 GlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu--- 239
    |||||
Db 442 GATGAGCAAAATTAGAG-----CATGCTAAACACGGCTTTAAAA 480
    :::::
Qy 240 ValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleArgAsnSerIle 259
    |||||
Db 481 ACCGCTTAGAGGGCGCAGGCTATT----- 507
    |||||
Qy 260 GlyGluValAlaValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAla 279
    |||||
Db 508 GGGAGCGTG----- 516
    |||||
Qy 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAsp----- 297
    |||||
Db 517 GTGGAGGTGGCCACAGAAAGGTCAAGTGGGTGCATTTATGATGCTGTTGATGTGAAT 576
    :::::
Qy 298 ---GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys 316
    ::::: |||||
Db 577 AGGGGGGATAGCATTTATATCAAAATCCATTTTATGAGGAAGCCGCAAAATTAAGACG 636
    :::::
Qy 317 AsnLeuIleGluAsnAlaSerAlaGluHis-----Gly 327
    ::::: |||||
Db 637 CGCATGATTGAATCTTTGAGTCGGAACAAAGCAAGACAGATTTTCATGGGTGGATGTGGGC 696
    :::::
Qy 328 TyrPheAspGlyArg----- 332
    |||||
Db 697 TTGAATGACGGAAATTCGGTTTAGATCAACTAGAAATAGGATTCTATCGGTATCCAAAGAT 756
    :::::
Qy 333 -----TrpLeuAspArgSerValAsp-----ValIleLeuProAsp 344
    ::::: |||||
Db 757 GTGTATATCGGTAGGGTTACTTAGACCGCTCATATTTCTCCCGCTTTTGTGAAACGGAT 816
    :::::
Qy 345 ---AsnThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGluTyrArgPhe 361
    ::::: |||||
Db 817 TTTTCTACCCATGACGCTTAAGCTTCATTATATAAGTCAAAAGAGGGGATCCAAATACAGGAT 876
    :::::
Qy 362 AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProasp 381
    ::::: |||||
```

```
Db 877 TCAGACATTTTA-----ATAGAGATTGACAACCCG 906
    :::::
Qy 382 LysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGluAla 401
    ::::: |||||
Db 907 GTAGTCCCTTAAAA-----ACCTTAGAAAAGCGCTTAAAGTGAAGAGAAAGATGTC 960
    :::::
Qy 402 TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu-----IleAlaThrArg 418
    ::::: |||||
Db 961 TTTAATATTGAGCATTTAAGAGCGGATCGCAAAATTTTAAAAACCGCAATCCCGCATAG 1020
    ::::: |||||
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
    ::::: |||||
Db 1021 GGTATTATGCTTGGCG-----GTGCTGAAGCCAGACTTGGATAAA- 1059
    :::::
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSer 458
    |||||
Db 1060 -----GATCAAAA 1068
    |||||
Qy 459 Thr---LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp--- 476
    |||||
Db 1069 AACGGGCTTGTCAAAGTCATTTATCGTATTGAAGTGGCGCATATGCTGTATATCAATGAT 1128
    :::::
Qy 477 -----IleSerProIleGluPheSerAlaSerAsnLeuIle- 488
    |||||
Db 1129 GTCATCATTTTCAGGAACCGGACGACGAGCATAGGATCATTTAGAGGGAGTTATTGTTA 1188
    :::::
Qy 489 -----GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
    :::::
Db 1189 GGGCTAAGGATATACTACAACTTGACCAACTGACA----- 1224
    :::::
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
    |||||
Db 1225 -----AATTCGCAAAATCTTTAAGCGCT--TTAGGATTC 1257
    :::::
Qy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
    |||||
Db 1258 TTCTCTAAAGTCAAAATTTGAAGAAAAAGGGTT-----AATAGTCA 1299
    :::::
Qy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
    :::::
Db 1300 CTCATGGATTTA----- 1311
    |||||
Qy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
    :::::
Db 1312 -----TTAGTACGCTAGAGAGGGCGCTACTGGGAG 1344
    :::::
Qy 587 Ile-----GlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGlu 604
    ::::: |||||
Db 1345 TTGCNAATTTGGGTAGGCTATGGCTCTTATGGAGGGCTTATGCTT-----AAT 1392
    :::::
Qy 605 HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAsp 624
    ::::: |||||
Db 1393 GGGAGCGTGAGCGAAAGAAACCTTTTGGCAGGCAAGCATGAGCTTGTATGCTAAC 1452
    :::::
Qy 625 -----LysLysGlyValLysLeuTyrAlaThrLysPro----- 635
    |||||
Db 1453 ATCGCTACAGGGGGGGGTAGATCTTATCCGGGCATGCCCAAAAGGAGCGGCGTATGTT 1512
    :::::
Qy 636 -----LeuSerProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
    |||||
Db 1513 GCCGGGAATTTGAGCTTGACTAATCCA----- 1539
    :::::
Qy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeu--- 669
    :::::
Db 1540 -----AGGATTTT-----CACAGCTGGTATAGCTCTACGATCAACCTTTAT 1581
    :::::
Qy 670 ---GluHisGluIleSerArgSerIleIleGlnAsnGlyGly----- 682
    ::::: |||||
Db 1582 GCGGATTACAGATAAGCTACCAATACATCCAAACAGCGGGGGCTTTGGGTGAATGTC 1641
    :::::
Qy 683 -----TrpAsnArgThrTyr---SerLeuArgTyrArgLeuAspLysLeuLys 697
    ::::: |||||
Db 1642 GGGCGCATGCTGGGGTAATAGAACCCCATGTGAGCTTAGGGTATAACTTGAATTTACCAAA 1701
    ::::: |||||
```

```
Qy 698 ThrGlnAlaProGluThrTrpGlnAspLeuProValAspPheValAsnGly----- 715
Db 1702 CTCCTTGTTTCAGCAGCCCTTTATACACCGCTACTATTCCTCTGTAAATGAAGTGGTT 1761
Qy 716 LysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAla----- 733
Db 1762 TCTCCAAGGAATGTTCTACCGCGCATCGTGATTATCAATCGCTTATCAGCGGTAATA 1821
Qy 734 -----AspAsnLeuValAsnPro-----Met 740
Db 1822 ACCCCCTTACAACCTGAAGCTGTTCTAGTCTGGAGCGGATCACCACCTTACCAGAAATA 1881
Qy 741 ArgGlyTyrArgGlnArg-----TyrSerLeuGluVal 751
Db 1882 AGAGGTATTTGGGATAGGGATTACCATACGCCATCACCAGCTCTTTTCACCCITGATGTG 1941
Qy 752 Gly-----SerSerGlyLeuValSerAspAlaAsn 761
Db 1942 AGCTATGACACACCGCATGATTATTACTTCCTAGAAATGGGTTATCTTTAGT----- 1995
Qy 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp----- 777
Db 1996 -----TCCTATGGCAGCATGCTGGCTTGCCAAAGCTCTGGCAGCGCTCAATCTTGG 2046
Qy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791
Db 2047 AACGGTTAGCGGGAATGCCGTAAACACCAAAAGTTTATGTGTAATTCGCCGCTTACCAC 2106
Qy 792 -----IleGlnAlaGlyTyr 796
Db 2107 CATTTGCAAAATATTATTATGATAGATTGTCGCTCGCTTTAAACGCAAGGAGGTAT 2166
Qy 797 IleTrpSerAspAsnPhe---AsnHisValProTyrArgLeuArgPheAlaGlyGly 815
Db 2167 ATCTTTAGGTATACACCGATGATTACTTGCCCTTAACCTCCACCTTCTATCGGGGGC 2226
Qy 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyr 835
Db 2227 GTAACCAACGGTGAAGCGTTTAGGACGGATCGGTTACTCTAAAGATGAGTTTGGCTTG 2286
Qy 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLys 855
Db 2287 TGGCTTGGAGCGATGGGATTTTACCGCTTCTACTGAATGAGTATGGGCTCTAAAG 2346
Qy 856 -----AspLeuArgLeuAlaValPheGlyAspIleGly----- 866
Db 2347 GCGCTAAATGCGCTTAGCGTGCTTTTGGACTTTGTTTCTTAACCTTTAAACCCCA 2406
Qy 867 -----Asn 867
Db 2407 ACTAGAGGGAGTTTTCCTATACGCTCTCTTTACGACAGCGAATTTTAAAGATTATGGC 2466
Qy 868 AlaTyrAspLysGlyPheThrAsnAspThr---LysIleGlyAlaGlyValGlyValArg 886
Db 2467 GTTATAGGGCTGGGTTTGAAGAGCCACTTGGAGGCGTCCACAGCGCTTCGACAGATTGAA 2526
Qy 887 TrpAlaSerProValGly 892
Db 2527 TGGATTTGCGCCATGGGG 2544
RESULT 35
AAV25138
ID AAV25138 standard; DNA: 2778 BP.
XX
AC AAV25138;
XX
DT 13-JUL-1998 (first entry)
XX
DE H. pylori cell envelope OMP ORF 05ep10815_4719175_c1_l15.
XX
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
```

```
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
KW cell envelope outer membrane protein; OMP; ds.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT CDS 1..2778
FT FT /*tag= a
XX
PN WO9737044-A1.
XX
PD 09-OCT-1997.
XX
PF 27-MAR-1997; 97WO-US05223.
XX
PR 06-DEC-1996; 96US-0761318.
PR 23-MAR-1996; 96US-0625811.
PR 02-APR-1996; 96US-0758731.
PR 25-OCT-1996; 96US-0736905.
PR 28-OCT-1996; 96US-0738859.
XX
PA (ASTR ) ASTRA AB.
XX
PI Alm RA, Smith D;
XX
DR WPI; 1997-503122/46.
XX
PT p-PSDB; AAW55729.
XX
PT Helicobacter pylori nucleic acid sequences and encoded
PT polypeptides(s) - useful in vaccines to treat or prevent H. pylori
PT infection and for diagnosis of H. pylori infection
XX
PS Claims 5.6.21; Page 929,930; 1145pp; English.
XX
XX This sequence encodes a H. pylori cell envelope outer membrane
XX protein having a terminal Phe residue.
XX The protein may be used in a vaccine to prevent or treat H. pylori
XX infection or to identify H. pylori life cycle activators or inhibitors. The
XX useful as potential H. pylori life cycle activators or inhibitors. The
XX DNA and probes derived from it may be used for the identification of
XX H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
XX acid sequences complementary to the DNA act as antisense sequences and
XX can be used to prevent the translation of H. pylori mRNA. Antibodies
XX against the protein can be used in immunoassays to evaluate the abundance
XX and distribution of H. pylori-specific antigens. The genomic sequence of
XX H. pylori (ATCC 55679) was determined from overlapping contigs generated
XX by mechanically shearing the bacterial DNA. The sequences were analysed
XX for ORF of at least 180 nucleotides, and the predicted coding regions
XX defined by computer evaluation. To identify likely H. pylori antigens for
XX vaccine development, the amino acid sequences predicted from various ORF
XX were analysed for significant homology to other known or exported
XX membrane proteins. Having identified and determined the sequences of
XX interest, particular regions can be isolated from H. pylori by PCR
XX amplification for recombinant polypeptide production, e.g. in E. coli
XX hosts.
SQ Sequence 2778 BP; 813 A; 514 C; 681 G; 770 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1 61e-06 Length: 2778
Score: 200.00 Matches: 206
Percent Similarity: 34.32% Conservatve: 143
Best Local Similarity: 20.26% Mismatches: 318
Query Match: 4.23% Indels: 352
DB: 18 Gaps: 51
```

US-09-914-168-2 (1-919) x AAV25138 (1-2778)

```
Qy 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnGlnIle 66
Db 128 ACGGCTCTAAACCAACG-----ATTTGGCTTCTCCAAAGAAACCC 169
Qy 67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
```

```
Db 170 CAAAAGAGCTCAAAAAATGAAGCTCAAAACGAACCT----- 208
Qy 87 ValAsnPheAspGlnSerProIleSerArgIleGly-----GluGlnSerProPro 104
Db 209 -----CTCAATCAATCAAAACGCCCTAAAGAAATGAAGTCAAGTCCATTT 253
Qy 105 LeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluLeuPhe 124
Db 254 CTTAT-GTCGGGCTTCTTAC-----ATGCTGACATGCTC 288
Qy 125 AlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGly 144
Db 289 GCTAATGAATTCGAAAGATTCGGTG-----GAC 318
Qy 145 GluGlnProAsnSerGluValValProThrLeuGluProGluLysProGlyLeu 164
Db 319 GATATGGTGGATTCAAAAAATAGACACCGCTGTTTACCTTGTCAACCAAGGTAT 378
Qy 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys 184
Db 379 TTTAAAGAGCTTTATGCC---ACTTTGAAACGGCATT-----TTAGAG 420
Qy 185 AlaLysPheTyrGlnSerSerGlnSerGly-----GluThrSerAlaIleGlySerSer 202
Db 421 TTTTCATTTTGTGAAACAGGAGTTCGGGGGTAGAAATCAAGGGTTATGGGACTGAA 480
Qy 203 HisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlu 222
Db 481 AAGGAAAAGACGGCTTAAATCCCAATGGGGATCAAAAAGGCCACACCTTTGATGAG 540
Qy 223 SerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu---ValAla 241
Db 541 CAAAAATTAGAG-----CATGCTAAAAACGGCTTTAAAAACGGCT 579
Qy 242 AlaArgAlaValClyTyrThrAspIleAspLeuSerIleIleArgAsnSerIleGlyGlu 261
Db 580 TTAGAGGGCAGGGCTTATTAT-----GGGAGC 606
Qy 262 ValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGlu 281
Db 607 GTG-----GTGGAG 615
Qy 282 ValArgGlyGluGlyAlaAspLysAlaPheThrThrValAlaAsp-----Glu 298
Db 616 GTGGCACAGAAAGGTCAGTGAGGAGCGCTTATTGATCGTGTGTTGATGTAATAGGGG 675
Qy 299 ValProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeu 318
Db 676 GACAGTATTTATCAACAATCCATTATGAGGGAAGCATATTAACACCCCGCTGTG 735
Qy 319 IleGluAsnAlaSerAlaGluHis-----GlyTyrPhe 329
Db 736 ATTGAATCTTTGACGCGCAACAGCAGCGGATTTTCATGGCTGGATGGGGCTTGAAT 795
Qy 330 AspGlyArg----- 332
Db 796 GACGGAAATTCGGCTTAGATCAATTAGAATACGATTTCTTGGTATCCAAGATGTGAT 855
Qy 333 -----TrpLeuAspArgSerValAsp-----ValIleLeuProAsp---Asn 345
Db 856 ATGCGTAGGGGTACTTAGACGCTCATATTTCTTCGCCCTTTTGTAAAAACGGATTTTCC 915
Qy 346 ThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGlnTyrArgPheAspGlu 363
Db 916 ACCCATGACGCTAAGCTTCATTATAGGTCAAGAGGGGATCCATACAGGATTTCCAGAT 975
Qy 364 ValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeu 383
Db 976 ATTTTA-----ATAGAGATTGACAAACCCGGTAGTC 1005
Qy 384 ProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsn 403
Db 1005 TTTTTTTT----- 1005
```

```
Db 1006 CCCTTAAAA-----ACTTAGAAAAAGCGCTTAAAGTTTAAAGGAAGATGTCTTTAAT 1059
Qy 404 LeuGlnAlaValArgAlaLeuSerAsnAspLeu-----IleAlaThrArgTyrPhe 420
Db 1060 ATTGACATTTTAAGACGGATCGCAAAATTTTAAAAACCGCAATCCCGCATTAAGGCTAT 1119
Qy 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
Db 1120 GCGTTTGGC-----GTGCTGAAGCCAGACTTGGATAAA----- 1152
Qy 441 SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThr--- 459
Db 1153 -----GACGAAAAAACCAGG 1167
Qy 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479
Db 1168 CTTGTGAAGATCATTTATCGTATTGAAGTGGCGAT-----ATGGTGCATCAATGAT 1221
Qy 480 IleGluPheSerAlaSerAsnLeuIleGlnLysLysLeuAsnLeuValAlaLysAla 499
Db 1222 GTCATCATTTTCAGGAAACCGCCAGCGCATAGGATC-----ATTAGG 1266
Qy 500 ArgHisLeuTyrAspMetProAspArg-----ValLeuAlaIleAsnHisAsp 516
Db 1267 AGGAATTTTACTAGGGCTTAAAGATAAATACAACCTTGACCAAACTGAGAAATCCGAA 1326
Qy 517 AspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArg 536
Db 1327 AATCTCTTCAGCGCT-----TTAGGTTTCTCTAAAGTCAAGATTGAAGAAAAAGG 1380
Qy 537 AlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeu 556
Db 1381 GTC-----AATAGCTCATTTGATGATTG----- 1404
Qy 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPhe 576
Db 1405 -----TTA 1407
Qy 577 ValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTyrGlySerAsp 594
Db 1408 GTGAGCGTAGAAGGGCGCCAGCGGCGAGTTGCAATTCGGTTGGCTATGGCTCTTAT 1467
Qy 595 ThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGln 614
Db 1468 GGAGGCTCATGCTT-----AATGGAGCGTGAGCGAAAGAAATCTTTTGGC 1515
Qy 615 AlaGlyAlaGluLeuArgLeuSerGluAsp-----LysLysGlyValLysLeuTyrAla 632
Db 1516 ACAGGCCAAGCATGAGCTTGTATGCTAACATTCGCCACAGGGGGGTAGATCTTATCCG 1575
Qy 633 ThrLysPro-----LeuSerHisProLeu 640
Db 1576 GCATGCCAAAAGGGCGGGCGGTATGTTTCCGGCAATTTGAGCTTGACTAATCCA--- 1632
Qy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsn 660
Db 1633 -----AGGATTTT-----GAC 1644
Qy 661 GlyPheAspLeuSerThrArgThrLeu-----GluHisLeuIleSerArgSerIleIle 678
Db 1645 AGCTGGTATAGCTCTACGATCAATCTTTATGCGGATTTACAGGATAAGCTACCAATACATC 1704
Qy 679 GlnAsnGlyGly-----TrpAsnArgThrTyr--- 687
Db 1705 CAACAAGGGCGGGGCTTTGGGGTGAATGTCGGCGCATCTGGGTAAATAGAACCCTATGTG 1764
Qy 688 SerLeuArgTyrArgLeuAsp-----LysLeu-LysThrGlnAlaProProGluThrTr 705
Db 1765 AGCTTAGGTATATAACTTGAATGTTACCAAACTCCTTTGGTTTCAGACGCCCTTATACAAC 1824
Qy 705 pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGln 725
Db 1825 CGCTACTATTCCTCTGTTAATGAAGTG-GCCTCTCCCAAGGCAATGTTCCACACCCGCATC 1883
```



Qy	223	SerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu---ValAla	241
Db	541	CAAAAATTAGG-----CATGCTAAACCGCTTTTAAAAACGGCT	579
Qy	242	AlaArgAlaValGlyTyrTyraSpIleAspLeuSerIleIleArgAsnSerIleGlyGlu	261
Db	580	TTAGAGGGCCAGGCTATTAT-----GGGAGC	606
Qy	262	ValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGlu	281
Db	607	GTG-----GTGGAG	615
Qy	282	ValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAsp-----Glu	298
Db	616	GTGCGACAGAAAGGTTCAGTCAGGAGCGGTATTGATCGTGTGTTGATGTCATAGGGGG	675
Qy	299	ValProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeu	318
Db	676	GACAGTATTATCAACAATCCATTATGAGGAGCGGATAAATTAACACCGCTGTG	735
Qy	319	IleGluAsnAlaSerAlaGluHis-----GlyTyrPhe	329
Db	736	ATTGAATCTTTGAGCGCGAACAAGCAGCGCATTTTCATGGCTGGATGTGGGGCTTGAAT	795
Qy	330	AspGlyArg-----	332
Db	796	GACGGGAAATTCGGCTTAGATCAATTAGAAATACGATCTTTTCGATATCCAAAGATGTGAT	855
Qy	333	-----TrpLeuAspArgSerValAsp-----ValIleLeuProAsp---Asn	345
Db	856	ATCGGTAGGGGTACTTAGCGCTCATATTCTTCGCCCTTTTGAACCGGATTTTTC	915
Qy	346	ThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGlnTyrArgPheAspGlu	363
Db	916	ACCATGACGCTAAGCTCCATTATAGGTCAAGAGGGGATCCATACAGGATTTTCAGAT	975
Qy	364	ValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeu	383
Db	976	ATTTTA-----ATAGAGATTGACACCCGGTAGTC	1005
Qy	384	ProValTysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsn	403
Db	1006	CCCTTAATA-----ACCTTAGAAAACGGCTTAAGTTAAAGAAAGAGATGCTCTTAAT	1059
Qy	404	LeuGlnAlaValArgAlaLeuSerAsnAspLeu-----IleAlaThrArgTyrPhe	420
Db	1060	ATTGACATTTAAGACGGATCGCGCAATTTTAAACACCGAAATCGCGATAGCGGTAT	1119
Qy	421	AsnMetValAsnThrGluLeValPheProGluArgGluGlnIleGlnAsnAspGlnVal	440
Db	1120	CGCTTTTTCG-----GTGTGAAGCCAGACTTTGGATAA-----	1152
Qy	441	SerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThr---	459
Db	1153	-----GACGAAAAAACCGG	1167
Qy	460	LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro	479
Db	1168	CTTCTCAAACTCATTTATCGTATTAGTGGCGCAT-----ATGGTGCATATCAATGAT	1221
Qy	480	IleGluPheSerAlaSerAsnLeuIleGlnAspTysLeuAsnLeuValAlaAlaLysAla	499
Db	1222	GTCCATCTTTCAGGGAACCGACGCGCAGCGCATAGGATC-----ATTAGG	1266
Qy	500	ArgHisLeuTyrAspMetProAspArg-----ValLeuAlaIleAsnHisAsp	516
Db	1267	AGGGAATTGTACTAGGGCTTAAGATAAATACAACCTTGACCAACTGAGAAATTCGCAA	1326
Qy	517	AspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArg	536
Db	1327	AATCTTTGAGCGCT-----TTAGGGTTTTCTCTAAAGTCAGATTGAAGAAAAAAGG	1380
Qy	537	AlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeu	556

[illegible]







polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (I) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AA564197-AA564364 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 2342 BP; 453 A; 738 C; 716 G; 435 T; 0 other;

**Alignment Scores:**

Pred. No.:	3.78e-05	Length:	2342
Score:	182.00	Matches:	139
Percent Similarity:	34.85%	Conservative:	91
Best Local Similarity:	21.06%	Mismatches:	243
Query Match:	3.85%	Indels:	188
DB:	23	Gaps:	25

US-09-914-168-2 (1-919) x AAS89745 (1-2342)

304	Qy	GlyAspValPheHisHis-----GlyLysTyrGluThrLysLysAsnLeu	318
361	Db	GGCGAGTTCCTCATCACTACAGCCACGAGCGAAAAAACAATGAATCAACCGTTA	420
319	Qy	IleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerVal	338
		:::	
421	Db	CTTTCCGTCAATAACCTGACCCACCTTTACGCCGGGCANA-----	462
339	Qy	AspValIleLeuProAspAsnThrAlaAspValSer-LeuIleTyrAspThrGlyThrG	358
		:::	
463	Db	-----GGCTTACGGCATGCTCTTTTGATTTATGCGCGGGG-----	499
358	Qy	nTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrTh	378
500	Db	-----AAGTGTGGGCATTTGTCGGG	519
378	Qy	rAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeu-----LeuTh	395
		:::	
520	Db	GAATCCG-----GCTCCGGG--AGACACCGCTCTGAAGTCGATCTCCGCGGCCTGAC	572
395	Qy	rValAsnMetGlyGluAla---TyrAsnLeuGlnAlaValArgAlaLeuSer-----	411
573	Db	GCCGACGAGGGGAAATTCACTACGAGAACCGTTCGCTGTATGCAATGAGCGAGGCCGA	632
412	Qy	----AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhePr	430
633	Db	CCGCGCGCGCTGCTGCGTACCGAAATGG---GGCGTGGTGCATCAGCATCCATCGACGG	689
430	Qy	oGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerSerArgTh	450
690	Db	CTCGCGCGCCAGGTCTCGCAGCGCGCAATATCGCGAGCGGCTGATGGCGACCGGGCG	749
450	Qy	rGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuTh	470
		:::	
750	Db	ACGTCATTCACGGCGATATTCGTCCCGCGCAGAGTGGCTGGAAGAGTGGAGATTCC	809
470	Qy	rAspGlyIleLeuMetAspLysSerProIleGluPheSerAlaSerAsnLeuIleGlnAs	490
		:::	
810	Db	CGCACACCGATCGACGACCTG---CCGACACCTTTTTCGCGCGT-----ATGCAGCA	860

Db 1749 CGGTGATTTTCGACAAAGTACCCGCGATCTGCGTTCTTCGCGGGGCGACCGCAGTAT 1808  
Qy 819 eArgGlyTyrAlaHisAsp-SerLeuSerProIleSerAspGly-TyrLeuThrGly 838  
Db 1809 TCGTGGCTACCAATACCAATCTATCGTCCGAAATACCCACGCTGACACTGAAAGG 1868  
Qy 839 GlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArg 858  
Db 1869 GCCTGAAGTTGATTAACCGGATCGTGGATACCACTACCACTGACCGGAAATGTACC 1928  
Qy 859 LeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLys 878  
Db 1929 AAGTGGCAGTTCGGGAA----- 1946  
Qy 879 IleGlyAlaGlyValArgTrpAlaSerProValGlyGlnValArgVal 896  
Db 1947 -----TCGCTGGCGCCC---GGCAGGTGGCGCTG 1973

RESULT 39  
AAS93088  
ID AAS93088 standard; cDNA; 1916 BP.  
AC AAS93088;  
DT 13-FEB-2002 (first entry)  
XX DNA encoding novel human diagnostic protein #28892.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.  
XX WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX PR 31-MAR-2000; 2000US-0540217.  
XX PR 23-AUG-2000; 2000US-0649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI: 2001-639362/73.  
XX P-PSDB; ABG28901.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
XX Claim 1: SEQ ID No 28892; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1916 BP; 349 A; 619 C; 582 G; 366 T; 0 other; 1

#### Alignment Scores:

Pred. No.:	3,51e-05	Length:	1916
Score:	181.00	Matches:	110
Percent Similarity:	34.00%	Conservative:	61
Best Local Similarity:	21.87%	Mismatches:	166
Query Match:	3.83%	Indels:	166
DB:	23	Gaps:	20

US-09-914-168-2 (1-919) x AAS93088 (1-1916)

Qy 458 SerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIle 477  
Db 349 GCCACCGCGCAGAAAGTGGCTGGAAGAGGTGGAGATTCCCGCAACCGGATCGACGACCTG 408  
Qy 478 SerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaA 497  
Db 409 ---CCGACACCTTTTCCGGCGGT-----ATCGACGACGCTTTCAGATT----- 450  
Qy 498 LysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAsp 517  
Db 451 ---GCCCGCAACCTGTGTGACGATCCG---AAGCTGTGTTTATGATGAACCGACCGCG 504  
Qy 518 GlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAla 537  
Db 505 GGGCTGGATGTCGTGGTGCGAGCGCCGCTGCTCGACCTGCTCGCGCGCTG----- 555  
Qy 538 IleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAla 557  
Db 556 -----GTGGTGGAGCTG-----AACCTCGCG 576  
Qy 558 AsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheVal 577  
Db 577 -----GTGGTGGATTGTA 588  
Qy 578 AlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrProGlySerAspThrGlyThr 597  
Db 589 ACCTCTTCCCGACGACGAGAC-----TGGGGCTTTTACCCACGAGCCT 630  
Qy 598 ArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAla 617  
Db 631 CGCTCGCGCGCGCTTC-----CCAGATCCACTCACTCGTCC 669  
Qy 618 GluLeuArgLeuSerGluAspLys----- 626  
Db 670 CGGAGCGCGCAACCGCGCGCCGCTATCCCGCGCGGAGCTGCCACACCCAGC 729  
Qy 627 -----GlyValLysLeuTyrAlaThrLysProLeuSer 637  
Db 730 GCACCCACAGTGGCGCGCGAGATGCGTCAAGCTC---GTCCCGCGCGCGCGCGC 786  
Qy 638 HisProLeuAsnAspGlnLeu-----ArgAlaThrLeuGlyTyrGlnGlnGlu--- 653  
Db 787 GTGACCCGAATGAGCGCGCGACCTTCCGACTCAGTAACTTTGGGGGTTCACACGAGCCC 846  
Qy 653 ----- 653  
Db 847 GGGGCGTGTATTGGGACCCCAATTCGACTTTTCCGACCTTCGACTTCAGCTATATAAT 906  
Qy 654 -----ValPheGlyHisSerThrAsnGlyPhe--- 662  
Db 907 GCGGTGCTGAAGAATCCACTGGACAATATTATTGGT-----GCAGCGCGTGTAAAG 960  
Qy 663 -----AspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIle 677  
Db 961 CGCACTGACCTGAACGATACCGAATCTGACTCCACTACGCTGGTGGCTTCTCGCTACTGG 1020

```
Qy 678 IleGlnAsnGlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLys 697
Db 1021 GATCTCTCCAGCGCTGCGAGCGTGCCATTAACTGCGCTGAGCTCTCGAC----- 1071
Qy 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValaspPheValAsnGlyLysPro 717
Db 1072 -----CACTTTACTCAGGGTGAAT 1092
Qy 718 SerGlnGluAlaLeuLeu-----AlaGlyValAlaValHisLysThrValAlaAspAsn 735
Db 1093 ACCAATACCAGTCGCTTTTATCTCGGGGTGATGATTACCCGCGAGCTTCCTCGTGT 1152
Qy 736 LeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGly 755
Db 1153 GGCCTGATGCCAACCTGGGCGGCTCGCAACGCTACTCTATCGACTACTCCAACACGCGC 1212
Qy 756 LeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe 775
Db 1213 TGGGGTTTCAGATGTCGATTCTCCGTTTTCAGCGCGAGAACGTCCTGGATCCGACACTG 1272
Qy 776 GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGly 795
Db 1273 TACGATCGC-----CATCGTTTGTTCACCGCGGCGACCTGGC 1311
Qy 796 TyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGly 815
Db 1312 TGGATTGAACCGGTGATTTCGACAAAGTACCGCGCGATCTGCTTCTTCGCGGGGCG 1371
Qy 816 AspGlnSerIleArgGlyTyrAlaHisAsp-SerLeuSerProIleSerAspLysGly-T 835
Db 1372 GACCGAGTATTCGTGGCTACCAATCAACATTCATCGGTCGCGAAATACCCACGCTGA 1431
Qy 835 yrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetL 855
Db 1432 CACTGAAAGGGCGCTCGAAGTGATACCGGATCGCTGGATACCAACACGTGACCG 1491
Qy 855 ysaspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrA 875
Db 1492 GAAATGTACCAAGTGCAGTTCGCGAA----- 1520
Qy 875 snaspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValA 895
Db 1521 -----TGGCTGGCGCCC---GGCGAGGTGC 1542
Qy 895 rgVal 896
Db 1543 GCCTG 1547

RESULT 40
AAQ66201
ID AAQ66201 standard; DNA; 2974 BP.
XX
AC AAQ66201;
XX
DT 08-DEC-1994 (first entry)
XX
DE H. influenzae SB3 D15 sequence.
XX
KW Vaccine; passive immunization; vector; antiserum; diagnosis; D15;
KW OMP; outer membrane protein; NTHi; ss.
XX
OS Haemophilus influenzae SB33.
XX
FH Key Location/Qualifiers
FT CDS 386..2764
FT /tag= a
XX
PN W09412641-A.
XX
PD 09-JUN-1994.
XX
PF 23-NOV-1993; 93WO-CA00501.
```

```
XX
PR 23-NOV-1992; 92GB-0024584.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Chong P, Klein M, Loosmore S, Sia DYC, Thomas W;
PI Yang Y;
XX
DR WP1: 1994-200269/24.
DR P-PSDB; AAR53757.
XX
XX Nucleic acid encoding D15 outer membrane protein - esp. of
PT Haemophilus influenzae, and related proteins, vectors, antisera
PT etc. useful in vaccines, for diagnosis and for passive
PT immunisation.
XX
XX Disclosure: Fig. 1D; 161pp; English.
XX
CC Outer membrane protein (OMP) D15 genes were isolated by screening
CC chromosomal libraries of H. influenzae type b (Hib) strains Ca,
CC Eagan and Minn A, and the non-typeable (NTHi) strains SB33 and PAK
CC 12085. Nucleotide sequences were determined for the D15 genes
CC (AAQ66198-202) and the corresponding aa sequences were derived
CC (AAR53734-58). D15 OMP can be produced easily and on a large scale,
CC free of other antigens and lipooligosaccharides, by recombinant DNA
CC methods using the isolated genes.
XX
SQ Sequence 2974 BP; 971 A; 471 C; 615 G; 917 T; 0 other;

Alignment Scores:
Pred. No.: 9,97e-05 Length: 2974
Score: 178.50 Matches: 173
Percent Similarity: 35.30% Conservative: 138
Best Local Similarity: 19.64% Mismatches: 345
Query Match: 3,78% Indels: 225
DB: 15 Gaps: 44

US-09-914-168-2 (1-919) x AAQ66201 (1-2974) 1
Qy 179 LysValProArgLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 374 AGATACAACTCGATGAAACAACTCTAATCGCAAGTTATTATTTCGTACGACAGCACT 433
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
Db 434 GTGTTTGGCGCA-----CCTTTTGGCAAAAGATATTTCGTGCGATGTT 478
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 479 GTTCAAGTGCATTTAGAACAAACAAATCCGAGCAAGTTTACCT----- 520
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 521 GTTCGTGCGCGTCAAGCGTGTGACTGACAAATGATGTG---GCTAATATTGTCGCTCTTTA 577
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 578 TTCGTAAGTGTGCTGATTCGATGATGTCAAAGCGCATCAAGAGCGCATGTGCTTGTGTT 637
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGly---AlaAsp 288
Db 638 AGCGTTGTGGCTAAATCGATCATTTTCAGATGTTTAAATCAAGGTAACTCTATTATTCCA 697
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 698 CCGAAGCACTAAACAAACAACTTAGATGCTAACGGGTTTAAAGTTGCGCATATTTTA--- 754
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 755 -----ATTCCAGAAAAATTAATGAATTTGCCAAAGTCTAAAGAGCACTATGCA 805
Qy 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
```

Db 806 AGTGAGTTCGCTAT---AAGCCAAACCGTTGAACCTATGTCATAACGCTACCAATAAT 862  
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
Db 863 CGCGCTGAATTTAAATCAATCAATCAAGATGATAAGCCAAATGGCATCAATTA 922  
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
Db 923 TTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTTACAAGAA 964  
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
Db 965 CAATGGGAATTCACACCTGATCTTGGTGGAAATATATGGGAAATAAATTAAGAGTGC 1024  
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 1025 CAATTCGAGAAAGATTTTCAGGCAATTCGT-----GATTAT 1060  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1061 TATTTAATAATATGGCTATGCCAAAGCACAAATCACTAAAGCGGATGTTTCAGCTAAATGAT 1120  
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456  
Db 1121 GAAAAAACAAGTTAATGTAAACCATGATGTAATGAAGGTTTACAGTATGACCTCGT 1180  
Qy 457 -----GluSerThrLeuGluProValIleGlu 465  
Db 1181 AGTCACGCATATATAGGTAATCTGGGAGGTATGTCGCCGAGCTTGAACCTTACTTTCA 1240  
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
Db 1241 GCATTACATTTAAATGATACTTTCGCCGCTAGTGATATTCGAGATGTAGAA----- 1291  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1292 ---AATGCAATTAAGCAAAACTTGGGAACGAGGTTACGGTAACACACAGATAATCT 1348  
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 1349 GTACCTGATTTGACGATGCAAAATAAACAATTTAGCGATAACCTTTGTTGTCATCGTA 1408  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1409 CGAGCTTTAACTGTTCCACCACTTCGCTTTGAAGCAAAATACCGTTTCTGCTGATAGTACT 1468  
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
Db 1469 TTACGTCAGGAATGCGGCCAACAGAGAACTTGGTATATATTCACATTTAGTTGAGTTA 1528  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1529 GGAATAATTCGCTTAGATCGTACAGGTTCTTCGAAACAGTTGAAACCGAATTCATCCT 1588  
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
Db 1589 ATCAATGCTAGCAATGATCAAGTGGATGCTGATATATAAGTCAAA----- 1633  
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592  
Db 1634 -----GACGTAACACGGGTAGTAGTCAACTTTGGTATGTTACGGT 1675  
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe-----GluHisAsnLeuIle----- 608  
Db 1676 ACAGAGAGTGTATTAGTTATCAACCAAGTGTCAACCAAGATAATTTCTTGGACACAGG 1735  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1736 CGGCAGCTAAGTATAGCTGGTACGAAAAATGATTATGTGTACGAGTGTCAATTTGGGTTAT 1795  
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1796 ACCGACCCCTATTTTACTAAAGAGTGTGAAGTCTTGTGGAAATGTTTCTTTTGGAAAC 1855

Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
Db 1856 TAGGATACTCTAAAGTCATACCTCTTAAGTATTAAGCGTACGACTTATGGAAGTAAT 1915  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis----- 671  
Db 1916 GTTACTTTAGGTTTCCCTGTAATAAGAAATAACTCTATTTATGATAGGATAGGCCATACC 1975  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686  
Db 1976 TATAATAAATTAGTAACCTTCTCTAGAA-----TATAACCGTAATTTATATATT 2026  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706  
Db 2027 CAATCAATGAAATTTAAAGGTAATGGCATTAACAACAAAT----- 2065  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726  
Db 2066 -----GACTTTGATTTTCTTTTGGTTGG----- 2089  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2090 -----AACTATAACAGCTTAATAGAGGCTATTTCCTCAACT 2125  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2126 AAAGGGGTTAAAGCAAGTCTTGGTGCAGAGTTACAATTCAGAGTCTCATACAATAAC 2185  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 2186 TACAACAACTAAGTCAGAGTGTACAGGTTTCTACCCATTA----- 2224  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2225 GACAGAGATCACTCTGGGTTGTATCTGCAAGAGCATCTGCAGGATATGCAATGGTTT 2284  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820  
Db 2285 GGAACAACAGCGTTTACCGTTCTCAAACTTATACAGCGGGTGGCATTTGTTTCATTACGC 2344  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2345 GCTTTTGTCTATGTTAGCATTTGGGCTTAACGCAATTTATCAAGGTCAAAATAATAAATTT 2404  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAla 847  
Db 2405 AATAAGATAAGTCTCTGAT-----GTGATGTTGGTGAATGCAATCGCTACAGCTACGCCA 2458  
Qy 848 Glu-----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAla 860  
Db 2459 GAGTTAAATGTGCCAACTCCATTTGTGAGTGATAAGAGTCAAAATACAGTCCGAACCTCC 2518  
Qy 861 ValPheGlyAspIleGlyAsnAlaTyr----- 869  
Db 2519 CTATTTGTGATGCGGCAAGTCTTTGGAATACTAAATGGAAATCAGATAAAATGGGATTA 2578  
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGly 880  
Db 2579 GAGACCAATGCTTGAAGACACTTACCGGATTTATGCAAAATCAAGCGGTACTCGCGCTCT 2638  
Qy 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900  
Db 2639 ACAGGTGTCGGATTCCAATGGCAATCTCTAGTAGGACAGTGTATTTTCTTTATGCTAAA 2698  
Qy 901 GlyValLys-----GluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrPro 918  
Db 2699 CCAATTAATAAATATGATAAATGATGTGCAACAGTTCCTCAATTTAGTATTGGGGTTCT 2758  
Qy 919 Phe 919  
Db 2759 TTC 2761

Tue May 6 09:37:17 2003

us-09-914-168-2.rng

Page 64

Search completed: May 6, 2003, 01:20:46  
Job time : 949 secs

1

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 5, 2003, 23:56:55 ; Search time 60 Seconds  
(without alignments)  
4697.266 Million cell updates/sec

Title: US-09-914-168-2

Perfect score: 4727

Sequence: 1 MSPXVLFANRSMFVALAAY.....TGKKEGNPKLHFFIGTFP 919

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=cgcn2\_1/USPTO.spool/US09914168/runat\_28042003\_151615\_6807/app\_query.fasta\_1.1095  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCLD=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09914168 @CGN\_1.1.32 @runat\_28042003\_151615\_6807 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA: \*  
1: /cgcn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgcn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgcn2\_6/ptodata/1/ina/6A.COMB.seq: \*  
4: /cgcn2\_6/ptodata/1/ina/6B.COMB.seq: \*  
5: /cgcn2\_6/ptodata/1/ina/PCTUS.COMB.seq: \*  
6: /cgcn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	181	3.8	2989	3	US-08-433-522A-9
2	181	3.8	2989	3	US-09-135-166-9
3	181	3.8	2989	4	US-08-942-046-9
4	178.5	3.8	2974	3	US-08-433-522A-7
5	178.5	3.8	2974	3	US-09-135-166-7
6	178.5	3.8	2974	4	US-08-942-046-7
7	172	3.6	2949	3	US-08-433-522A-1
8	172	3.6	2949	3	US-09-135-166-1
9	172	3.6	2949	4	US-08-942-046-1
10	172	3.6	2950	3	US-08-433-522A-5
11	172	3.6	2950	3	US-09-135-166-5
12	172	3.6	2950	4	US-08-942-046-5

13	172	3.6	2984	3	US-08-433-522A-3	Sequence 3, Appli
14	172	3.6	2984	3	US-09-135-166-3	Sequence 3, Appli
15	172	3.6	2984	4	US-08-942-046-3	Sequence 3, Appli
16	172	3.6	2987	3	US-08-433-522A-55	Sequence 55, Appl
17	172	3.6	2987	3	US-09-135-166-55	Sequence 55, Appl
18	172	3.6	2987	4	US-08-942-046-55	Sequence 55, Appl
19	159.5	3.4	3957	1	US-07-689-008-5	Sequence 5, Appli
20	159.5	3.4	9540	1	US-07-689-008-1	Sequence 1, Appli
21	137.5	2.9	13121	4	US-08-961-527-126	Sequence 126, App
22	136.5	2.9	38584	4	US-09-453-702B-50	Sequence 50, Appl
23	136	2.9	5976	4	US-08-621-944A-2	Sequence 2, Appli
24	136	2.9	5976	4	US-08-945-567D-2	Sequence 2, Appli
25	136	2.9	6973	1	US-08-478-370-1	Sequence 1, Appli
26	136	2.9	6973	4	US-08-483-855C-1	Sequence 1, Appli
27	136	2.9	6973	4	US-08-621-944A-1	Sequence 1, Appli
28	136	2.9	6973	4	US-08-945-567D-1	Sequence 1, Appli
29	136	2.9	6973	4	US-08-431-718C-1	Sequence 1, Appli
30	136	2.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
31	135	2.9	5629	4	US-09-453-702B-243	Sequence 243, App
32	135	2.9	30549	4	US-09-134-001C-322	Sequence 322, App
33	134.5	2.8	9377	4	US-09-221-017B-1002	Sequence 1002, Ap
34	134	2.8	3121	3	US-08-961-083-117	Sequence 117, App
35	132.5	2.8	32207	2	US-08-770-379-20	Sequence 20, Appl
36	132.5	2.8	32207	4	US-08-757-669A-20	Sequence 20, Appl
37	132.5	2.8	32207	4	US-09-230-371A-20	Sequence 20, Appl
38	131.5	2.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
39	131	2.8	11091	4	US-09-134-001C-2243	Sequence 2243, Ap
40	130.5	2.8	16836	4	US-09-147-236-1	Sequence 1, Appli
41	130.5	2.8	16836	4	US-09-147-236-10	Sequence 10, Appl
42	130	2.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
43	129	2.7	3758	3	US-08-323-477-1	Sequence 1, Appli
44	129	2.7	14736	4	US-08-961-527-171	Sequence 171, App
45	128.5	2.7	2028	4	US-09-134-001C-1710	Sequence 1710, Ap

#### ALIGNMENTS

##### RESULT 1

US-08-433-522A-9  
; Sequence 9, Application US/08433522A  
; Patent No. 6013514  
; GENERAL INFORMATION:  
; APPLICANT: CHONG, Pele  
; APPLICANT: THOMAS, Wayne  
; APPLICANT: YANG, Yan Ping  
; APPLICANT: LOOSMORE, Sheena  
; APPLICANT: SIA, Dwo Yuan Charles  
; APPLICANT: KLEIN, Michel  
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6TH Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/433-522A  
; FILING DATE: 12-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163

## ; INFORMATION FOR SEQ ID NO: 9:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 390..2768
; US-08-433-522A-9
```

## Alignment Scores:

```
Pred. No.: 2,11e-09 Length: 2989
Score: 181.00 Matches: 172
Percent Similarity: 35.57% Conservative: 141
Best Local Similarity: 19.55% Mismatches: 345
Query Match: 3.83% Indels: 222
DB: 3 Gaps: 44
```

US-09-914-168-2 (1-919) x US-08-433-522A-9 (1-2989)

```
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 378 AGGATCAATCGATGGAAGAACTCTAATCGCAAGTTTATTTCGGTGGGACACCGACT 437
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 438 GTGTTTCGCCGA-----CCCTTTGTGCCAAAGATATTTCGTGTGGATGGT 482
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 483 GTTCAAGGTGACTTAGAACACAAATCCGAGCAAGTTTACCT-----524
Qy 239 LeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 525 GTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 582 TTCGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 642 AGCCTGTGGCTAAATCGATCATTTTCAGATGTTTAAATCAAAGGTAACTCTGTTATTCCC 701
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 702 ACTAAGCAGCTTAACAAACTTAGATGCTAACCGGTTTAAAGTTGGCGATGTTTA--- 758
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 759 -----ATTCCAGAAAAATTAAATGAATTTGCCAAAAGTGTAAAAGAGCAGCATATGCA 809
Qy 329 PheAspGlyArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 810 AGTGTAGTCTGCTAT---AAGCGAACCGTTGAACCTATTGTCAATACGCTGCCAAATAAT 866
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 867 CGTCTGNAATTTTAATTCAAATCAATGAAGATGATTAAGCAAAATTTGGCGATCACTAACT 926
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 927 TTC-----AAGGGGAACGAA---TCTGTGTAGTAGCAGGTACATTTACAAGAA 968
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr-----402
Db 969 CAAATGGAATTTACAACTGATTTCTTGTGGAAATTTATGGGGAAATAAATTTGAAGGTGCG 1028
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 1029 CAATTCCGAGAAGATCTGCAGCAATTCGT-----GATTAT 1064
```

```
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1065 TATTTAAATAATGCTATGCCAAGACACAAATCACTAAACAGGATGTTTCAGCTAAATGAT 1124
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp-----456
Db 1125 GAAAAACAAAGTTAATTAACCATTTGATGCTAATTAAGGTTTACAGTATGACCTTCGT 1184
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1185 AGTGCACGCATTTATAGTAATCTCGGAGGTATGTCGCCGAGCTTGAACCTTTACTTTCA 1244
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1245 GCATTACATTTAATATGATACTTCCCGCGTAGTGATATTGCGAGATGTAGAA-----1295
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1296 ---AATCAATTTAAAGCAAACTTTGGGAAACGAGGTTCACGGTAACACACAGTAAATCT 1352
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1353 GTACCTGATTTTGACGATGCAAAATAAACATTAACGATAACCTTTGTTGTTGCTGCTGA 1412
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1413 CGAGCTTTAACTGTTGCCCAACTTTCGCTTTGAAGGAATAATCTCTGCTGATGATGACT 1472
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1473 TTACGTGAGGAATGCGCAACAAGAGAACTTGGTGATATAATTCACAAATTTAGTTGAGTTA 1532
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1533 GGAATAATTCGCTTAGATCGTACAGGTTTCTCGAAACAGTTTGAACCAAGTAAATGATCCT 1592
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1593 ATCAATGGTAGCAATGATGAAGTGGATGCTCGTATATAAAGTCAAA-----1637
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1638 -----GAACGTTAACCGGTAGTAGTATCAACTTTGGTATTGGTTACGGT 1679
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle-----608
Db 1680 ACACAGAGTGGTATCAGTTATCAACAAGATATTAAACAAGATAATTTCTTGGGAACAGGG 1739
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1740 GCGCAGCTAAGTATAGCTGTGTACGAAAAATGATTATGTCAGAGTGTCAATTTGGGTAT 1799
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1800 ACCGAACCTATTTTACTAAAGATGGGTAGTCTCTGGTGAATATTTCTTTGCAAAAC 1859
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1860 TACGATAACTCTAAAGTGATACATCCTCTAACTATAAGCGTACGACTTATGGAAGTAAT 1919
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis---671
Db 1920 GTTACTTAGTTCCTGTAATCAAAATAACTCCTATTATGTAGGATTAGGCCATACC 1979
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr---686
Db 1980 TATTAATTAATTAATTAATTTGCTCTCTAGAA-----TATAACCGTAATTTATATATT 2030
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuThrGlnAlaProGluThrTrpGln 706
Db 2031 CAATCAATGAATTTAAGGTAATGGCATTAAACAAAT-----2069
```



```
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2070 -----GACTTTGATTTTCTTTGGTTGG----- 2093
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyTyrArgGlnArg 746
Db 2094 -----AATTATAACAGCCTTAATAGAGCGCTATTTCCCAACT 2129
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2130 AAGGGGTTAAAGCAAGTCTTGGTGGAGAGTACTATTCAGGTTCTGATAACAATAC 2189
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2190 TACAACTAAGTCAGATGATACAGGGTTCTACCACTA----- 2228
Qy 783 AsnArgAlaHisGln-----MetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2229 GACAGAGATCACCGCTGGTGTGATCTGCAAAAGCATCTGCGAGGTGCAAAATGGTTT 2288
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2289 GGAACAAGCGTTTACCGTTCTATCAACTTATACAGCGGTGGCATTTGTCATTAGC 2348
Qy 821 GlyTyrAlaHisAspSerLeuSerPro-----IleSerAspLysGlyTyrLeuThr 837
Db 2349 GGTTCGTTATGTTAGTATTGGGCTTAATGCAATTTATGCGAATGTTGTAATGGTACT 2408
Qy 838 -----GlyGlyGlnValLeuAlaValGlyThrAlaGlu 848
Db 2409 TTTAATAGATAAGTCTCGATGTGTTGTAATGCAATCAACACTGCGAGTGCAGAA 2468
Qy 849 -----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAlaVal 861
Db 2469 CTTATTGTACCACTCCATTTGTGAGTGATAAAGCCAAATACAGTCCGAACCTCCCTA 2528
Qy 862 PheGlyAspIleGlyAsnAlaTyr----- 869
Db 2529 TTTTCTGATGCGCAAGTGTGTTGTAATGCAATGCAATGCAATGCAATGCAATGCAAT 2588
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGlyAla 881
Db 2589 AGCAAGGTCTTGAAGACTTACCTGATTTATGCAAAATCAACCGTATTCGGCGCTCTACA 2648
Qy 882 GlyValGlyValArgTyrPalaSerProValGlyGlnValArgValAspValAlaThrGly 901
Db 2649 GGTGTCGATTCCAATGCAATCTCTATGACCATTTGTTATTTCTTATGCTAAACCA 2708
Qy 902 ValLys-----GluGluGlyAsnProIleLysLeuHisPheIleGlyThrProPhe 919
Db 2709 ATTAATAAATATGAAATGATGATGTCGAACAGTTTCCAAATTTAGTATGGGGGCTCTTTC 2768

RESULT 2
US-09-135-166-9
; Sequence 9, Application US/09135166
; Patent No. 6083743
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,166
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 390..2768
; US-09-135-166-9

Alignment Scores:
Pred. No.: 2,11e-09 Length: 2989
Score: 181.00 Matches: 172
Percent Similarity: 35.57% Conservative: 141
Best Local Similarity: 19.55% Mismatches: 345
Query Match: 3.83% Indels: 222
DB: 3 Gaps: 44

US-09-914-168-2 (1-919) x US-09-135-166-9 (1-2989)
```

```
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 378 AGGATACAATCGATCGAAACAACTTCTAATCGCAAGTTTATTATTCGGTGGCACAACGACT 437
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 438 GTGTTGCCGCA-----CCTTTTGCCAAAGATATTCTGTTGGATGGT 482
Qy 219 IleThrGlnGluSerAlaMetAspLysSerIleProArgLeuArgGlnThrAla 238
Db 483 GTTCAAGGTGACTTAGAACAAACAAATCGAGCAAGTTTACCT----- 524
Qy 239 LeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleleArgAsnSer 258
Db 525 GTTCGTGCTGTCGCGGTGACTGACATGATG---GCTAATATTGTCGCTCTTTA 581
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 582 TTCGTAAGTGGTCGATTCGATGTGTAAGCGCATCAAGAGCGATGCTGCTGTTGTT 641
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 642 AGCGTTGTCGCTAAATCGATCATTCAGATGCTTAAATCAAGGTAACCTCTGTTATTC 701
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 702 ACTGAAGCACTTAAACAAACAACTTAGATGCTAACCGGGTTTAAAGTTGGCGATGTTT 758
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 759 -----ATTGAGAGAAAAATTAATGAAATTTCGCAAAAGTGTAAAGAGACACTATGCA 809
```



```

Sequence 9, Application US/08942046
Patent No. 6264954
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: THOMAS, Wayne
APPLICANT: YANG, Yan Ping
APPLICANT: LOOSMORE, Sheena
APPLICANT: SIA, Dwo yuan Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/942,046
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/433,522
FILING DATE: 12-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2989 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 390..2768
US-08-942-046-9

```

```

Alignment Scores:
Pred. No.: 2,11e-09 Length: 2989
Score: 181.00 Matches: 172
Percent Similarity: 35.57% Conservative: 141
Best Local Similarity: 19.55% Mismatches: 345
Query Match: 3.83% Indels: 222
DB: 4 Gaps: 44

```

```

US-09-914-168-2 (1-919) x US-08-942-046-9 (1-2989)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 378 AGGATCAATCGATGAAGAAATCTCTAATCGCAAGTTTATTATTCGGTGGCGCAACGACT 437
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLysAlaLeuGluAsp 218
Db 438 GTGTTTGGCCGA-----CCTTTTGTCCAAAGATATTCGTGTGGATCGT 482
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 483 GTTCAAGGTGACTTAGAACACAAATCCGAGCAAGTTTACCT----- 524
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258

```

```

Db 525 GTTCGTCTGTGTGACGCGTGACGTGACAATGATGTG---GCTAATATTGTCCGCTCTTTA 581
QY 259 Ile-----GlyGluVal---AspValIleIleHisLeuGlyGluProValTyrIle 275
Db 582 TTCGTAAGTGGTCGATTCGATGATGTTGAAAGCGCATCAAGAAAGCGATGTGTTGTTGTT 641
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 642 AGCGTTGCTGCTAAATCGATCATTTTCAGATGTTAAATCAAAAGGTAACCTCTGTTATTCCC 701
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 702 ACTGAAGCACTTAAACAAACCTTAGATGCTAACGGGTTTAAAGTTGGCGATGTTTAA--- 758
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 759 -----ATTCAGAGAAATTAATGAATTTGCCAAAAGTGTAAAAGACGACATATGCA 809
QY 329 PheAspGlyArgTirLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 810 AGTGTAGTGTGCTAT---AACGCAACGTTTGAACCTATTGTCAATACGCTGCCAAATAAT 866
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 867 CGTGTGAAATTTTAAATTCAAATCAATGAAGATGATAAGCAAAATTTGGCATCATTAAC 926
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 927 TTC-----AAGGGGACGAA---TCTGTTAGTAGCAGTACATTTACAAGAA 968
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 969 CAAATGGAATTTACCAACCTGATTCTTGTGTGAAATTTATGGGGAATAATTTGAAGTGG 1028
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 1029 CAATTCGAGAAAGATCTGAGGCAATTCGT-----GATTAT 1064
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1065 TATTTAAATATGCTATGCCAAAGCACAAATCACTAAACGGGATGTTTCAGCTAATGAT 1124
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1125 GAAAAACAAAAAGTTAATGTAAACCATTCATGTAATGAAGCTTTACAGATATGACCTTCGT 1184
QY 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1185 AGTCACGCAATTATAGTAATCTGGGAGGTATGCTCGGAGCTTGAACCTTTACTTTCA 1244
QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1245 GCATTACATTTTAAATGATATCTTCGCGCGTAGTGTATATTCAGATGTAGAA----- 1295
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLysAlaLysAlaArgHisLeuTyrAsp 504
Db 1296 ---AATGCAATTAAGCAAAATCTTGGGACGAGGTACCGGTAAACACACAACTAATCT 1352
QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1353 GTACCTGATTTTGACGATGCAAAATAAACATTAGCGATAACCTTTGTTGTTGATCGCTGGA 1412
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1413 CGAGCTTTAACTGTTCGCCAATCTCGCTTTTGAAGGAATATACCGTTCTGCTGATAGTACT 1472
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1473 TTACGTGAGGAATGCGCAACAAAGAGGAACCTTGGTATAATTAATCAATTAATTTAGTTG 1532
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562

```

```

Qy 862 PheGlyAspIleGlyAsnAlaTyr----- 866
    ||| ||| :::: ::
Db 2529 TTTGTTGATCGGCAAGTGTTCGGAATACTAAATGCAAAATCAGATAAAATGCATTAGAG 2588
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGlyAla 881
    ||| ||| :::: ::
Db 2589 AGCAAGTCTTTGAAGACTTTACTGTATTATGGCAAAATCAAGCGGTATTGCGCGCTCTACA 2648
Qy 882 GlyValGlyValArgTTPAlaSerProValGlyGlnValArgValAlaIleThrGly 901
    ||| ||| :::: ::
Db 2649 GGTGCGGATCCANTGGCAATCTCCTATTGGACCATTTGTTTCTTATGCTAAACCA 2708
Qy 902 ValLys-----GluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
    ::||| ||| :::: ::
Db 2709 ATTAATAAAATATGAATATGATGATGTCGAACACAGTTCCAATTTAGTATTGGGGCGCTCTTTC 2768

RESULT 4
US-08-433-522A-7
: Sequence 7, Application US/08433522A
: Patent No. 6013514
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: THOMAS, Wayne
: APPLICANT: YANG, Yan Ping
: APPLICANT: LOOSMORE, Sheena
: APPLICANT: SIA, Dwo Yuan Charles
: APPLICANT: KLEIN, Michael
: TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Slim & McBurney
: STREET: 6TH Floor, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/433,522A
: FILING DATE: 12-SEP-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2974 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 386..2761
: US-08-433-522A-7

```

US-09-914-168-2 (1-919) x US-08-433-522A-7 (1-2974)

Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198  
Db 374 AGGATACAATCGCAAAACCTTCTAATCGCAAGTTTATTATTCGGTACGACAGACT 433  
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218  
Db 434 GTGTTTCCGCA-----CCTTTGTGGCAAAAGATATTCGTGTGATGGT 478  
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238  
Db 479 GTTCAAGTGACTTAGAACACAATCCGAGCAAGTTTACCT----- 520  
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258  
Db 521 GTTCGTGCGGTGACGGTGTGACTGACAAATGATG-----GCTAATATTGTCGGTCTTTA 577  
Qy 259 Ile-----GlyGluVal--AspValIleIleHisAspLeuGlyGluProValTyrIle 275  
Db 578 TTCGTAAGTGTGCGATTGCGATGTGAAAGCGCATCAAGGCGGATGCTTGTGTT 637  
Qy 276 AspTyrArgAla-----ValGluValAlaArgGlyGluGly---AlaAsp 288  
Db 638 AGCGTTGTGCTAATCGCATTCATGATGTTAAATCAAGGTAACCTCTATTATTCGA 697  
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308  
Db 698 CCTGAAGCACTAAAACAACTTAGATGCTAAACGGGTTAAAGTTGGCGATATTTA--- 754  
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328  
Db 755 -----ATTCGAGAAAATTAATGAATTTGCCAAAGTGTAAAAGCACATATCA 805  
Qy 329 PheAspGlyArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsn 345  
Db 806 AGTGTAGTCCGTAT--AACGCAACCGTTGACCTATTGTCAATACCGCTACCAATAAT 862  
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
Db 863 CGCGCTGAATTTTAAATTCAAATCAATGAAGATGATAAGCAAAATGGCATCATTA 922  
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrValAsnMetGlyGluAlaTyr--- 402  
Db 923 TTC-----AAGGGAACGAA---TCTGTTAGTACGAGTACATTACAGAA 964  
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr--- 1024  
Db 965 CAATGGAAATTACAACCTGATCTCTGGTGGAAATTTATCGGGAATAAATTTCAAGGTCCG 1060  
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 1025 CAATTCCAGAAAGATTTCCAGGCAATTCGT-----CATTAT 1060  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1061 TATTAAATAATGCGTATGCCAAAGCACAAATCACTAAAGCGGATGTTTCAGCTAAATGAT 1120  
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456  
Db 1121 GAAAAACAAGTAATGTAAACCATTTGATGTAATGAAGGTTTACAGTATGACCTTCGT 1180  
Qy 457 -----GluSerThrLeuGluProValIleGlu 465  
Db 1181 AGTGCACCATATAGGTAATCTGGGAGGTATGTCGCGGAGCTTGAACCTTTACTTTCA 1240  
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
Db 1241 GCATTACATTTAAATGATACCTTTCGCGCGTAGTGATATTCAGATGTAGAA----- 1291  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1292 ---AATGCAATTAAGCAAAACCTTGGGGAACGAGGTTTACGGTAAACACACAGTAATTTCT 1348

Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspGlyValAsn 520  
Db 1349 GTACCTGATTTTGACGATCAATNAACATTAGCGATAACCTTTCTGTTGATGCTGGA 1408  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1409 CGACGTTTAACTGTTTCCACCACTTCGCTTTGAAGGAATAACGCTTCTGCTGATAGTACT 1468  
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
Db 1469 TTACGTGAGAAATCGCCCAACAAGAAGAACTTGTGTATAATTTCACAAATTAGTTGAGTTA 1528  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1529 GGAAAAATTCGCTTAGATCGTACAGGTTTCTCGAAACAGCTTGAAACCGCAATTTGATCCT 1588  
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
Db 1589 ATCAATGCTAGCAATGATGAAGTGGATGCTGATATAATAAAGTCAAA----- 1633  
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGluIle-----GlyLeuGlyTyrGly 592  
Db 1634 -----GAACGTAAACCGGTAGTAGTCACTTTGCTATTGTTACGGT 1675  
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608  
Db 1676 ACAGAGAGTGTATTAGTTATCAAGCAAGTGTCAACAAGATAAATTTCTTGGCAACAGGG 1735  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1736 GCGGCAGTAAGTATAGCTGGTACGAAATAATGATTATGGTACGAGTGTCAATTTGGGTTAT 1795  
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1796 ACCGAGCCCTATTACTAAAGATGCTGTAAGTCTTGGTGGAAATGTTTCTTTGAAAC 1855  
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
Db 1856 TACGATAACTCTAAAGTATACATCCTCTAACTATTAAGCGTACGACTTATGGAAGTAAT 1915  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
Db 1916 GTTACTTTTAGGTTTCCCTGTAATAAGAAATAACTCTATATGTAGTAGTATAGGCCATACC 1975  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr----- 686  
Db 1976 TATAATAAAATTAGTAACCTTCTCTAGAA-----TATAACCGTAATTTATATATT 2026  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706  
Db 2027 CAATCAATGAAATTTAAAGGTAATGGCAATTAACAACAAAT----- 2065  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726  
Db 2066 -----GACTTTGATTTTCTTTTGGTGG----- 2089  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2090 -----AACTATAACAGCCTTAATAAGAGGCTATTTTCCCACT 2125  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2126 AAAGGGTTAAAGCAAGTCTTGTGGCAGGATTACAATTCAGGTTCTGATAACAATAC 2185  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 2186 TACAACAACTAAGTCAGATGATACAGGTTCTACCCCATTA----- 2224  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2225 GACAGAGATCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGTTT 2284

```
QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2285 GGAACAAGCGTTTACGGTTCTATCAAACTTATACACGGGTGGCATTTGGTTCAATACGC 2344
QY 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2345 GGTTTTCCTTATCGTACGATGGCGCTAACGCAATTTATCAAGGTCAAAATAATAAATTT 2404
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGlyThrAla 847
Db 2405 AATAAGATAAGTCTCAT-----GTGATTGGTGGTAATGCAATCGCTACAGCTACGGCA 2458
QY 848 Glu-----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAla 860
Db 2459 GAGTTAATTGTGCAACTCCATTGTGTGAGTGATAAGAGTCAAAATACAGTCCGAACCTCC 2518
QY 861 ValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2519 CTATTTGTTGATCGGCAAGTGTGTTGGAATCTAAATCGAATCAAAATGGAATGGAATTA 2578
QY 870 -----AspLysGlyPheThrAsnAspThrLysIleGly 880
Db 2579 GAGAGCAATGCTCTGAAGACTTACCGGATTATGCGCAAAATCAAGCGGTACTCGCGCTCT 2638
QY 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900
Db 2639 ACAGTGTGCGGATTCGAATGGCAATCTCCTAGTGGACCACTGTTATTTCTTATGCTAAA 2698
QY 901 GlyValLys-----GluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrPro 918
Db 2699 CCAATTAAAAAATATGAAAATGATGATGTCGAACAGTTCCAATTTAGTATGGGGTTCT 2758
QY 919 Phe 919
Db 2759 TTC 2761

RESULT 5
US-09-135-166-7
Sequence 7, Application US/09135166
Patent No. 6083743
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: THOMAS, Wayne
APPLICANT: YANG, Yan Ping
APPLICANT: LOOSMORE, Sheena
APPLICANT: SIA, Dwo Yuan Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6TH Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/135,166
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/433,522
FILING DATE: 12-SEP-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb
```

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2974 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURES:
NAME/KEY: CDS
LOCATION: 386..2761
US-09-135-166-7

Alignment Scores:
Pred. No.: 3,95e-09 Length: 2974
Score: 178.50 Matches: 173
Percent Similarity: 35.30% Conservative: 138
Best Local Similarity: 19.64% Mismatches: 345
Query Match: 3.78% Indels: 225
DB: Gaps: 44

US-09-914-168-2 (1-919) x US-09-135-166-7 (1-2974)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 374 AGGATACAACTCGATGATAAAGAACTTCTAATCGCAAGTTTATTTCGGTACGACACGACT 433
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 434 GTGTTTGGCGCA-----CCTTTGTGGCAAAAGATATTTCGTGTGGATGT 478
QY 219 IleThrGlnGlnLysAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 479 GTTCAAGGTGACTTAGAACACAAATCCGACCAAGTTTACCT----- 520
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 521 GTTCGTGCGCGGTGACGTGACATGATGATG--GCTAATAATTGTCGCTCTTTA 577
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 578 TTCGTAAGTGGTGGATTCGATGATGAAAGCGCATCAAGAGCGCATGTCTTGTGTT 637
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGly---AlaAsp 288
Db 638 AGCGTTGTGCGCTAAATCGATCATTTTCAGATGTTAAATCAAAAGTAACTCTATTATCCA 697
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 698 CCGTAAGCACTAAACAAACAACTTAGCTAAACGGGTTTAAAGTTGGCGATATTTTA--- 754
QY 309 HisGlyLysTyrGluThrLysLysAlaLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 755 -----ATTTCGACAAAAATTAATGAATTTGCCCAAGTGTAAAGAGCACTATGCA 805
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 806 AGTGTAGGTGCGTAT---AACGCAACCGTTGAACCTATTGTCAAATACGCTACCAATAAT 862
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 863 CGCGCTGAAATTTTAAATTCAAATCAATGAAGATGATAAGCCCAAAATTCGATCATTAAC 922
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 923 TTC-----AAGGGGAACGAA---TCTGTAGTAGCAGTACATTACAAGAA 964
QY 386 LysArgGluLeuGluGlnLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 965 CAAATGGAATTACAACTCATCTCTTCTGCGAAATTTATGCGGAAATAAATTTGAAGGTGCG 1024
```

Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 1025 CAATTCGAGAAAGATTTTCAGGCAATTCG-----GATTAT 1060  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1061 TATTTAAATAATGCGTATGCGCAAGCACAATCACTAAGCGGATGTTTCAGCTAAATGAT 1120  
Qy 439 GlnValSerPheGluGlnSerSerArgThrGlnProAlaGlnValAsp-----456  
Db 1121 GAAAAACAAAAGTTAATGTAACCATTCATGTAATGAAGGTTTACACGATGACCTTCGT 1180  
Qy 457 -----GluSerThrLeuGlnProValIleGlu 465  
Db 1181 AGTCACGCATATAGGTAATCTGGAGGTATGTCGCGAGCTTGAACCTTTACTTTCA 1240  
Qy 466 ThrValGluLeuThrAspGlyLeu---LeuMetAspIleSerProIleGluPheSerAla 484  
Db 1241 GCATTACATTTAAATGATACCTTCGCGCGTAGTGATATTCAGATGTAGAA-----1291  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1292 --AATGCAATTAAGCAAACTTTGGGCAACGAGGTTACGGTTAACACACAGTAATTC 1348  
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspGlyValAsn 520  
Db 1349 GTACCTGATTTTGACGATGCAAAATAAAACATTTAGCGATAACCTTTGTTGATCTGGA 1408  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1409 CGAGCTTTAACTGTTCCACCACTTCGCTTTGAAGGAATACCGTTTCGCTGATAGTACT 1468  
Qy 534 ValAlaArgAlaIleLeuProAspIleSer-----GluAsnGluValIleAspLeu 550  
Db 1469 TTACGTCAGGAATCGCGCAACAAGAGAACTTGGTATAATTCACAATTAGTTGAGTTA 1528  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1529 GGAATAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTTGAAACCGAATTCATCCT 1588  
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
Db 1589 ATCAATGCTAGCAATGATGAAGTGGATGCTGTATATAAGTCAAA-----1633  
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592  
Db 1634 -----GAACGTAACACACGCGGTAGTATCAACTTTGGTATTGGTTACGGT 1675  
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle-----608  
Db 1676 ACAGAGAGTGTATAGTTATCAAGCAAGTGTCAACCAAGATAATTTCTTGGGAACAGGG 1735  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1736 CCGCAGTAAAGTATAGTCTGCTACGNAATATGATGTCAGGTCATCAATTTGGGTTAT 1795  
Qy 622 SerGlu-----AspLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1796 ACCGAGCCCTATTTTACTAAAGTGTGTAAGTCTTGGTGAATAATGTTTCTTTGAAAC 1855  
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
Db 1856 TAGGATACTCTAAAGGTGATACATCCCTCACTAATAAGCGTACGACTATGGAAGTAAT 1915  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis---671  
Db 1916 GTTACTTTAGGTTTCCCTGTAATGCAAAATACTCTATTATGATAGGATAGGCCATACC 1975  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr-----686  
Db 1976 TATAATAAAATAGTAACCTTGCTCTAGAA-----TATAACCGTAATTTATATATT 2026  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706

Db 2027 CAATCAATGAAATTTAAAGGTAATGCGATTAACAAAT-----2065  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726  
Db 2066 -----GACTTTGATTTTCTTTTGTGG-----2089  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2090 -----AACTATAACAGCCTTAATAGAGCGCTATTTCCTCCAACT 2125  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2126 AAAGGGGTTAAAGCAAGTCTTGGTGCAGAGTTACAATTCAGGTTCTCATACAAATAC 2185  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 2186 TACAACACTAAGTGCAGATGTACAGGTTTCTACCCATTA-----2224  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2225 GACAGAGATCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCAATGCTTT 2284  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820  
Db 2285 GGAACAACGCGTTTACCGTTCTTCAAACTTATACAGCGGTTGGCATTTGTTTCATACGC 2344  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro-----829  
Db 2345 GGTTCGTTTGTATGTTAGTACATTCGGCTACGCAATTTATCAAGGTCAAAATAATAATTT 2404  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAla 847  
Db 2405 AATAAGATAAGTTCTCAT-----GTGATTGGTGAATGCAATCGCTACAGCTAGCGCA 2458  
Qy 848 Glu-----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAla 860  
Db 2459 GAGTTAATGTCGCAACTCCATTTGTGAGTGATAGAGTCAAAATACAGTCCGCACTCC 2518  
Qy 861 ValPheGlyAspIleGlyAsnAlaTyr-----869  
Db 2519 CTATTTGTTGATGCGCGCAAGTCTTGGATACTAATAATGGAATCAGATAAAATGGATTA 2578  
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGly 880  
Db 2579 GAGAGCAATGCTCTGAAAGACTTACCGGATTTGCGAAATCAAGCGCTACTCGCGCTCT 2638  
Qy 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900  
Db 2639 ACAGGTGTCGGATTCGAATGGCAATCTCTAGTGGACCACTGATTTCTTTATGTCTAAA 2698  
Qy 901 GlyValLys-----GluGluGlyAsnProIleLysLeuHisPheIleGlyThrPro 918  
Db 2699 CCAATTAAATAATATGAAATGATGATGATGTCGCAACAGTTTCAATTTAGTATTGGGGTTCT 2758  
Qy 919 Phe 919  
Db 2759 TTC 2761

## RESULT 6

US-08-942-046-7  
; Sequence 7, Application US/08942046  
; Patent No. 6264954  
; GENERAL INFORMATION:  
; APPLICANT: CHONG, Pele  
; APPLICANT: THOMAS, Wayne  
; APPLICANT: YANG, Yan Ping  
; APPLICANT: LOOSMORE, Sheena  
; APPLICANT: SIA, Dwo Yuan Charles  
; APPLICANT: KLEIN, Michel  
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney  
STREET: 6TH Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/942,046  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/433,522  
FILING DATE: 12-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-732 MIS-jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2974 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 385..2761  
08-942-046-7

Alignment Scores:	3.95e-09	Length:	2974
red. No.:	178.50	Matches:	173
Score:	35.30%	Conservative:	138
Percent Similarity:	19.64%	Mismatches:	345
Self Local Similarity:	3.78%	Indels:	225
Very Match:	4	Gaps:	44

US-09-914-168-2 (1-919) x US-08-942-046-7 (1-2974)

```

179  LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
374  AGGATAACATCGATTGCAAAAACCTTCTAATCGCAAGTTTATTATTCGTACGCAACGACT 433
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
199  IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434  GTGTTGTCGGCA-----CCITTTGTGGCAAAAGATATTCTGTGGATGTT 478
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
219  IleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
479  GTTCAAGTGACTTAGAACACAAATCCGACCAAGTTTACCT----- 520
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
239  LeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleIleArgAsnSer 258
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
521  GTTCGTGCCGTGACGCTGATGACATGATGTG--GCTAATATTGTCCGCTCTTTTA 577
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
259  Ile-----GlyGluVal---AspValIleIleHisAspLeuGluProValTyrIle 275
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
578  TTCGTAAGTGGTGCATGTCGATGATGTGAAAGCCGATCAAGAAGCGCATGTGTTGTGTT 637
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
276  AspTyrArgAla-----ValGluValArgGlyGluGly---AlaAsp 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
638  AGCGTTGTGGCTAAATCGATCATTTTCAGATGTTTAAATCAAGAGTAACCTATTATTCCA 697
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
289  AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308

```



Db 1676 ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTCAACACAGATAAATTTCTTGGGAACAGGG 1735  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1736 GCGGCAGTAAGTATAGCTGGTACGAAATGATTGGTACGAGTGTCAATTTGGGTAT 1795  
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1796 ACCGAGCCCTATTATTACTAAGATGCTGAAGTCTTGGTGAATGTTTCTTTGAAAC 1855  
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSer 658  
Db 1856 TAGATAACTTAAAGTAGTACATCTCTAATAAGCGTACGACTTATGGAAGTAAT 1915  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
Db 1916 GTTACTTTAGGTTTCCTCTAATAAGTAATACTCTATTATGTAGGATTAGCCATACC 1975  
Qy 672 -----GluLeuSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr----- 686  
Db 1976 TATAATAAATTAGTAACCTTGGTCTAGAA-----TATAACCGTAATTATATATT 2026  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706  
Db 2027 CAATCAATGAATTTAAGGTATGCGATTAAACAAAT----- 2065  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuAlaGlyVal 726  
Db 2066 -----GACTTTGATTTTCTTTGGTGG----- 2089  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2090 -----AACTATAACAGCCTTAATAGAGGCTATTTCCCAACT 2125  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2126 AAGGGGTTAAGCAAGTCTTGGTGACGAGTTACAATTCAGGTTCTGATAACAATAC 2185  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 2186 TACAAACTAAGTCAGATACAGGCTTCTACCCATTA----- 2224  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2225 GACAGAGATCACCTCTGGTGTCTATCTGCAAAAGCATCTGCAAGATATCAATGGTTTT 2284  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820  
Db 2285 GGAACAACGCTTTACCGTTCTATCAAACTTATACAGCGGGTGGCATTTGTTTCATACGC 2344  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2345 GGTTTTGTCTTATGTAGCATTTGGCTTAACGCAATTTATCAAGGTCAAAATAATAAATTT 2404  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGlyThrAla 847  
Db 2405 AATAAGATAAGTCTGAT-----GTGATTGGTGGTAAATGCAATCGCTACAGTCAGCGCA 2458  
Qy 848 Glu-----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAla 860  
Db 2459 GAGTTAATTTGTCCCAACTCCATTTGTGAGTGATAGAGTCAAAATACAGTCCGAACTCC 2518  
Qy 861 ValPheGlyAspIleGlyAsnAlaTyr----- 869  
Db 2519 CTATTGTGTGATGCGCAAGTGTTCGTAATACTAAATGGAATCAGATAAAATGCGATT 2578  
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGly 880  
Db 2579 GAGAGCAATGCTCTTGAAGACTTACCGGATTATGCGCAATCAAGCGGTACTCGCGCTCT 2638  
Qy 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900  
Db 2639 ACAGGTGTCGGATTCCAAATGGCAATCTCTAGTGGACCATGTTATTTCTTTATGCTAAA 2698

Qy 901 GlyValLys-----GluGluGlyAsnProIleLysLeuHisPheIleGlyThrPro 918  
Db 2699 CCAATTAATAAATATGAATATGATGTGCAACAGTTCCAATTTAGTATTGGGGTTCT 2758  
Qy 919 Phe 919  
Db 2759 TTC 2761  
RESULT 7  
US-08-433-522A-1  
; Sequence 1, Application US/08433522A  
; Patent No. 6013514  
; GENERAL INFORMATION:  
; APPLICANT: CHONG, Pele  
; APPLICANT: THOMAS, Wayne  
; APPLICANT: YANG, Yan Ping  
; APPLICANT: LOOSMORE, Sheena  
; APPLICANT: SIA, Dwo Yuan Charles  
; APPLICANT: KLEIN, Michel  
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6TH Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/433,522A  
; FILING DATE: 12-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2949 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 75..2465  
; US-08-433-522A-1  
Alignment Scores:  
Pred. No.: 2,03e-08 Length: 2949  
Score: 172.00 Matches: 171  
Percent Similarity: 35.21% Conservative: 141  
Best Local Similarity: 19.30% Mismatches: 344  
Query Match: 3.64% Indels: 230  
DB: Gaps: 44  
US-09-914-168-2 (1-919) x US-08-433-522A-1 (1-2949)  
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198  
Db 63 AGCATACAATCGATGAAAAAATCTCTAATCGCAAGTTTATTATTCGTTACGACACGACT 122  
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218  
Db 199 SerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218

Db 123 GTGTTTCCCGCA-----CCTTTTGGCAAAAGATATTCGTGGTGGT 167  
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238  
Db 168 GTTCAAGGTGACTTAGAACACAAATCCGAGCAAGTTACT----- 209  
Qy 239 LeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleIleArgAsnSer 258  
Db 210 GTTCGGTCCGGTCAGCGTGTGACTGACAATGATGTG---GCTAATATGTCCTCTTTA 266  
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275  
Db 267 TTCGTAAGTGGTCGATTCGATGATGTGAAGCCATCAAGAGCGGATGTGCTTGTGT 326  
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288  
Db 327 AGCGTTGTGCTAAATCGCATTCATTTAGATGTTAAATCAAAAGTAACCTGTTATCC 386  
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308  
Db 387 ACTGAAGCACTTAACAAACACTTAGTCTAACGGGTTTAAAGTTGGCGATGTTTAA 443  
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328  
Db 444 -----ATTCGAGAAAAATTAATGAATTTGCCAAAGTGTAAAGAGCACTATGCA 494  
Qy 329 PheAspGlyArgTrpLeuAspArgSerValaspValIle-----LeuProAspAsn 345  
Db 495 AGTGAGTGGCGTAT---AACCCACAGTGTGAACCTATTGTCAATACGCTACCAATAAT 551  
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
Db 552 CGCGTGAAATTTAATCAATCAATGAGATGATGAAGCAAAATTTGGCATTCATAACT 611  
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
Db 612 TPTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTACAGAA 653  
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
Db 654 CAATGGAAATACACACCTGATCTGGTGGAAATATATGGGAAATATAATTTGAAGTCCG 713  
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 714 CAATTCGAGAAAGATTTCCAGTCAATTCGT-----GATTAT 749  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 750 TATTTAAATAATGGTATGCCAAAGCACAAATTTACTAAACCGGATGTTTCAGCTAAATGAT 809  
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456  
Db 810 GAAAAACAAAAGTTAATGTAAACCATTCATGTAATGAAGGTTTACAGTATGACCTCGT 869  
Qy 457 -----GluSerThrLeuGluProValIleGlu 465  
Db 870 AGTCACGCAATTATAGTAATCTGGAGGTATGCTCGGAGCTTGAACCTTTACTTTCA 929  
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
Db 930 GCATTACATTTAATGATCTTCCGCGGTAGTGATATTGCAGATGTAGAA----- 980  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 981 ---AATGCAATTAAGCAAACTTTGGAGACCGCGTTACGGTAGCGCAACGGTAAATTC 1037  
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 1038 GTACCTGATTTTGATGATGCAAAATAAAACATTACCGATAACCCCTTGTGTGTGCTGGA 1097  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1098 CGAGGTTTAACGTGTCGCAACATTCGCTTTGAAGGAATAACCGTTCTGCTGATGACCT 1157

Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
Db 1158 TTACGTCAGGAAATGCCCAACAAGAAGAACTTGGTATAAATTCACAATTAGTTGAGTTA 1217  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1218 GGAAAAATTCGCTTAGTCTACAGGTTCTTCGAAACAGTCCGAAACCGAATGATGCTC 1277  
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
Db 1278 ATCAATGGTAGTAATGATCAAGTGGATGTCGTATATAAAGTCAAA----- 1322  
Qy 575 ValPheValAlaSerAspLysProArgaspGlyGlnIle-----GlyLeuGlyTrpGly 592  
Db 1323 -----GAACGTAAACAGCGGTAGTATCAACTTTGGTATTCGTTACGGT 1364  
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608  
Db 1365 ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTAAACAAGATAATTTCTTGGCAACAGGG 1424  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1425 GCGGCAGTAACTATAGCTGGTACGAAATAATGATTATGGTACGAGTCTCAATTTGGTTAT 1484  
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1485 ACCGAGCCCTATTTTACTAAAGATGGTGAAGTCTTGGTGGAAATGTTTTCTTTGAAAAAC 1544  
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
Db 1545 TACGATAACTCTAAAGTGATACATCCCTAACTATAAGCGTACGACTTACGGAAGTAA 1604  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
Db 1605 GTTACTTTAGGTTTCCCTGTAATGAAATAAATAACTCCTATTATGTAGGATTAGGTACATACC 1664  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686  
Db 1665 TATAATAAATAGTAACCTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 1715  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706  
Db 1716 CAATCAATGAAATTTAAAGGTAATGGCATTAACAAATA----- 1754  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726  
Db 1755 -----GACTTTGATTTTCTTTTGGTTGG----- 1778  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyTyrArgGlnArg 746  
Db 1779 -----AACTATAACAGCTTAAATACAGGCTATTTCACCAACT 1814  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 1815 AAAGGGGTTAAAGCAAGCTCTGGTGACGAGTACTACTATCCAGGTTCTGTGAACAAATAC 1874  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 1875 TACAAACTAAGTSCAGATGTACAGGTTTCTACCCATT----- 1913  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 1914 GACAGAGATCACCTCTCTGGTGTGATCTGCAAAAGCATCTCGAGGATATGCAATGGTTTT 1973  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820  
Db 1974 GGAACACAGCGTTTACCGTCTCTATCAAACTTATACAGCGCGGTGGCATCGGTTTCATTAC 2033  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2034 GGTTTTGTGTTATGTTAGTATTGGACCAATTTATGCCCAATATGTTGTAATGTGACT 2093

```
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlnValLeu 842
Db 2094 GGTACTGGTACTTTTAAGAGATAAGTTCTGAT-----GTGATTGGTGGTAATGCAATC 2147
QY 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2148 GCTACAGCTAGCGCAGAGTAATTTGGCCAACCTCCATTTGTGAGCGATAAGAGCCAAAT 2207
QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2208 ACGGTCGCGACCTCTCTATTGTTGATCGGCAAGTGTGGAACTAAATGGAATCA 2267
QY 870 -----AspLysGlyPheThrAsn 875
Db 2268 GATAAAATGGATTAGAGCGCATGATTAAAAAGATTCCCTGATTATGCCAAATCAAGC 2327
QY 876 AspThrLysIleGlyAlaGlyValAlaArgTTPAlaSerProValGlyGlnValArg 895
Db 2328 CQTATTGCGGCTCTACAGGTGTGGATTCCCAATGGCAATCCTATTGGGCCATTGGTA 2387
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2388 TTCCTCTATGCCCAACCAATTAATAAATATGAAATGATGATGTCGAACAGTTCCTCAATTT 2447
QY 914 PheIleGlyThrProPhe 919
Db 2448 AGTATTGGAGTTCTTTC 2465

RESULT 8
US-09-135-166-1
; Sequence 1, Application US/09135166
; Patent No. 6083743
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN.
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,166
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2949 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 75..2465
US-09-135-166-1

Alignment Scores:
Pred. No.: 2,03e-08 Length: 2949
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 3 Gaps: 44

US-09-914-168-2 (1-919) x US-09-135-166-1 (1-2949)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 63 AGGATACAATCGATGATAAAAACTTCTAATCGCAAGTTTATTATTTCGGTAGCAACGACT 122
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
Db 123 GTGTTGCCGCA-----CCTTTGTGCAAAAGATATTCTGTGGATGCT 167
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 168 GTTCAAGGTGACTTAGACAACAATAATCCGAGCAAGTTTACCT----- 209
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 210 GTTCGTGCGCGTCAGCGTGTGACTGACAATGATGTG---GCTAATATTGTCCGCTCTTTA 266
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 267 TTCGTAAGTGTGTCGATTCGATGTGTAAGCCGCTCAAGAGCGATGCTGTGTTGTT 326
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 327 AGCGTTGTGCTAAATCGATCATTTTCAGATGTTTAAAAATCAAGGTAACCTCTGTATTCCC 386
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 387 ACTGAAGCACTTAAACAAACTTAGATGCTAACGGGTTTAAAGTTGGCGATGTTTA--- 443
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 444 -----ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGTAAAAGAGCACTATGCA 494
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 495 AGTGTAGGTGCGTAT---AACGCAACAGTTGAACCTATTGTCAATACGCTACCAAAATAAT 551
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 552 CGCGCTGAATTTTAATCAATCAATGAAGATGATGAACCAAAATGGCATCATTAACCT 611
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 612 TTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGCTACATTACAAGAA 653
QY 386 LysArgGluLeuLeuGluGlnLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 654 CAAATGGAATTTACAACCTGATTCTTGTGGAATTTATGGGAAATTAATTTGAAGGTGCG 713
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 714 CAATTCGAGAAAGATTTCGAGTCAATTCGT-----GATTAT 749
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 750 TATTTAAATAATGGCTATGCCAAAGCACAAATTAATAAACGGATGTTTACAGCTAAATGAT 809
```



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/942,046

Alignment Scores:		
Pred. No.:	2.03e-08	Length:
Score:	172.00	Matches:
Percent Similarity:	35.21%	Conservative:
Best Local Similarity:	19.30%	Mismatches:
Query Match:	3.64%	Indels:
DB:	4	Gaps:
		44
		2949

US-09-914-168-2 (1-919) x US-08-942-046-1 (1-2949)

Qy	LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGluThrSerAla	198
	:::::	
Db	AGGATAAATTCGATGAAGAACAATTCAATCGCAAGTTTATTATTCGGTAGCACAACGACT	122
	:::::	
Qy	IleGlySerSerHisGlnIysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp	218
	:::::	
Db	CTTTTGCCGCA-----CCTTTTGTGCCAAACATATTCTGTGGATGCT	167
	:::::	
Qy	IleThrGlnGluSerAlaMetAspLeuAnSglySerIleProArgLeuArgGlnThrAla	238
	:::::	
Db	GTTCAAGGTGCATTAGAACCAACAATCCGAGCAAGTTTACCT-----	209
	:::::	
Qy	LeuValAlaAlaArgAlaValGlyTyrTyrrAspIleAspLeuSerIleIleArgAsnSer	258
	:::::	
Db	GTTCCGCGCGTCAGCGTGTGACTGCATCATGATGTC---GCTAATATTGTCGCCTCTTTA	266
	:::::	
Qy	Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle	275
	:::::	
Db	TTCGTAAAGTCGGTCGATTCGATGATGTGAAGCCGATCAAGAGCGCATGTGCTGTGTGTT	326
	:::::	
Qy	AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp	288
	:::::	
Db	AGCTTTGTGGCTAAATCGATCATTTTCAGATGTTTAATAACAAGTAACCTGTTATTCCTC	386
	:::::	
Qy	AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis	308
	:::::	
Db	ACTGAACACTTAAACAAAACTTTAGATGCTAACGGGCTTTAAAGTTGCGCATGTTTA---	443
	:::::	
Qy	HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr	328
	:::::	
Db	ATTCCAGAAAATTTAAATGAATTTGCCAAAAGTGTAAAGAGCACATATGCA	494
	:::::	
Qy	PheAspGlyValqTrpLeuAspArgSerValAspValIle-----LeuProAspAsn	345

```
Db 1485 ACCGAGCCCTATTCTTACTAAAGATGGTGTAAAGTCTTGGTGAAATGTTTCTTTGAAAC 1544
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1545 TACGATAACTCTAAAGATGATACATCTCTAACTATAGCGTAGGACTTACGGAAGTAAT 1604
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1605 GTTACTTAGGTTTCCCTGTAATGAAATGAAATAAATCTCTATTATGATGAGATTAGGTATACC 1664
Qy 672 ---GluLeuSerArgSerIleGlnAsnGlyClyTrpAsnArgThr----- 686
Db 1665 TATAATAAATAGTAACTTGTCTAGAA-----TATAACCGTAAATTTATATATT 1715
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706
Db 1716 CAATCAATGAAATTAAGGTAATGATGCAATTAACAAT----- 1754
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726
Db 1755 -----GACTTTGATTTTCTTTTGGTTG----- 1778
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 1779 -----AACTATAACAGCCTTAATAGAGGCTATTTCCTCAACT 1814
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 1815 AAAGGGGTTAAAGCAAGTCTTGGTGACGAGTACTATTCCAGGTTCTGATAACAAATAC 1874
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 1875 TACAACTAAGTCAGATGATCAGAGGTTCTACCCATTA----- 1913
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 1914 GACAGATACACTCTGGTTGTATCTGCAAAGCATCTGCAGGATATGCAATGGTTT 1973
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArg 820
Db 1974 GGAAACAAGCGTTTACCGTTCTATCAAACTTATACAGCGGGTGCATCGGTTCAATAGCT 2033
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2034 GCTTTGCTTATGTTAGTATTGGACCTACCGCAATTTATGCCGAATATGGTAATGGTAGT 2093
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2094 GGTAAGTGTACTTTTAAAGATAGTTCAT-----GTGATTGGTGGTAATGCAATC 2147
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2148 CTTACAGCTAGCGGAGAGTTAATTTGTGCCAACTCCATTTGTGAGCGATAAGAGCCAAAT 2207
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2208 ACGGTCCGAACCTCTTATTGTTGATGCGCGAAGTGTGGAACTACTAAATGAAATCA 2267
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2268 GATAAAAATGGATPAGAGCGGATGATTAAAAAGATTGCTGATTATGCAAAATCAAGC 2327
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2328 CGTATTCGCCCTCTACAGGTGCGGATTCCAATGGCAATPCTCCTATTGGGCATTTGGTA 2387
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2388 TTCTCTTATGCCAAACCAATTAATAAATATGAAATATGATGATGTCGACAGTTCCTCAATTT 2447
Qy 914 PheIleGlyThrProPhe 919
Db 2448 AGTATTGGAGTCTTTC 2465
```

```
RESULT 10
US-08-433-522A-5
; Sequence 5, Application US/08433522A
; Patent No. 6013514
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN 1
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1K7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/433,522A
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-434 MLS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 5:
; LENGTH: 2950 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 334..2724
; US-08-433-522A-5

Alignment Scores:
Pred. No.: 2,03e-08 Length: 2950
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 3 Gaps: 44

US-09-914-168-2 (1-919) x US-08-433-522A-5 (1-2950)
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyClyThrSerAla 198
Db 322 AGGATACATCGATGAAAAAATCTTAATCCAGGTTTATTATTCGGTACGACAAACGACT 381
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
Db 382 GTGTTTGGCCGCA-----CCTTTTGTGGCAAAAGATATTCGTGTGCGTGTGT 426
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 427 GTTCAAGGTGACTTAGAACACAAATCCGACGAAATTTACCT----- 468
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 468
```

Db 469 GTTCGTGGCGGTACCGTGTGACTGACAAATGATGTG---GCTAATAATTGTCGCCCTCTTTA 525  
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275  
Db 526 TTCGTAAGTGGTCGATTCGATGATGTGAAGCGCATCAAGAGCGGATGTCTGTGTGTT 585  
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288  
Db 586 AGCGTTGTGGCTAAATCGATCATTCAGATGTTAAATCAAAGTAACTCTGTATTTCCTCC 645  
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308  
Db 646 ACTGAAGCACTAAACAAACTTAGACTAAGCGGTTTAAAGTTGCGCATGTTTAA--- 702  
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328  
Db 703 -----ATTCGAGAAAAATTAATGAATTGCCAAAAGTCTAAAGCAGCACTATGCA 753  
Qy 329 PheAspGlyArgTrpLeuAspArgSerValaspValIle-----LeuProAspAsn 345  
Db 754 ACTGTAGGTTCGGTAT---AACGCAACAGTTGAACCTATTGTCAATACGCTACCAATAAT 810  
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
Db 811 CCGCTGAAATTTTAATCCAAATCAATGAAGATGATAAGCAAAATGCGCATCAITTAAC 870  
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
Db 871 TTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTACAAGAA 912  
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
Db 913 CAAATGGAATTACAACCTGATTCCTGGTGAAATATATGGGAAATAAATTTGAAGGTGCG 972  
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 973 CAATTGCGAAGAGATTGCGAGTCAATTCGT-----GATTAT 1008  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1009 TATTTAAATTAATGGCTATGCCAACACACAAATTAATAAACGGATGTTTCAGCTAATGAT 1068  
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValasp----- 456  
Db 1069 GAAAAACAAAGTTAATGTAACCATTTGATGTAATGAAGTTTACAGTATGACCTTCGT 1128  
Qy 457 -----GluSerThrLeuGluProValIleGlu 465  
Db 1129 AGTGCACGCATTAATAGGTAACTGGAGGTATGCTGCGCAGCTTGAACCTTTACTTTCA 1188  
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
Db 1189 GCATTACATTTTAATGATACTTTCGCCCGCTAGTGATATTTGCAGATGTAGAA----- 1239  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1240 ---AATGCAATTAAGCAAAACTTGAGAACCGGTTACGGTAGCGCAACGGTAAATTTCA 1296  
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 1297 GTACCTGATTTGATGATGCAATAAACAATAGCCGATACCCCTGTTGTTGATGCTGGA 1356  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1357 CGACGTTTAAGCTTCGCCAACTTCGCTTTGAAGGAATAATACCGTTTCTGCTGATAGCACT 1416  
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
Db 1417 TTACGTCAGGAAATGCGCAACAAAGGAAGTAATGTTGATATAATTCACAATTAGTTGAGTTA 1476  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1477 GGAATAATTCGCTTAGCTACAGGTTTCTTCGAAACAGTCGAAACCGAATTTGATCCT 1536

Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
Db 1537 ATCAATGGTAGTAATGATGAAGTGGATGCTGATATATAAAGTCAAA----- 1581  
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592  
Db 1582 -----GAACGTAACACGGTAGTAGTACCACTTTGTTGTTACGGT 1623  
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608  
Db 1624 ACAGAGAGTGTATTAGTTATCAAGCAAGTGTAAACAAGATRAATTTCTTGGGAACAGGG 1683  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1684 GCGCAGTAGTATAGTCTGTCACGAAAAATGATTATGGTACGAGTGCAATTTGGGTAT 1743  
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1744 ACGAGCCCTATTTTACTAAAGATGGTGAAGTCTTGGTGGAAATGTTTCTTTTGAANAAC 1803  
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
Db 1804 TACGATAACTCTAAAAGTAGTACATCTCTAACTATTAAGCGTAGGACTTAGGGAAGTAAT 1863  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
Db 1864 GTTACTTTAGGTTTCCTCTGTAATGAAAAATAACTCTATTATGTAGGATTTAGGTCATACC 1923  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686  
Db 1924 TATAATAAAATAGTAACCTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 1974  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706  
Db 1975 CAATCAATGAATTTAAGGTAATGGCATTAACAACAAAT----- 2013  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726  
Db 2014 -----GACTTTCATTTTCTTTGTTGG----- 2037  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2038 -----AACTATACAGCCTTAATAGAGGCTATTTCCCAACT 2073  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2074 AAAGGGTTAAACACAGTCTTGTGGACGAGTTACTATTCAGGTTCTGTATAACAATAAC 2133  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 2134 TACAAACTAAGTCAGATGTACAGGGTTTCTACCCATTA----- 2172  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2173 GACAGAGATCACTCTCGGTTGTATCTGCAAAAGCATCTCGCAGATATGCAAAATGGTTT 2232  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820  
Db 2233 GGAACAACGGTTTACCGTTCTATCAAACTTATACACGGGTGGCATCGTTTCAATACGT 2292  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2293 GGTTTTCCTATGTTAGTATTTGGACCTAACGCAATTTATGCCGAATATGTTAATGGTAGT 2352  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842  
Db 2353 GGTACTGGTACTTTAAGAAGATAAGTCTTGAT-----GTGATTTGGTGAATGCAATC 2406  
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856  
Db 2407 GCTACAGTAGCCGACAGTTAATTTGGCCAACCTATTTGTGACCGATAAGAGCAAAAT 2466

```

QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2467 ACGGTCCGAACCTCTTATTGTGATCGCGCAAGTGTGGAATACTAAATGGAATCA 2526
QY 870 -----AspLysGlyPheThrAsn 875
Db 2527 GATAAAATGGATTAGACGGATGATTAAAGATTGGCTGATTATGGCAATCAAGC 2586
QY 876 AspThrLysIleGlyAlaGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2587 CGTATTCGCCCTCTACAGGTGCGGATTCGAATGCAATCTCTATTGGCCATGGTA 2646
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2647 TTCTCTATGCCAAACCAATAAATAATGAAATGATGATGCGAACAGTTCCAATTT 2706
QY 914 PheIleGlyThrProPhe 919
Db 2707 AGTATTGGAGGTTCTTTC 2724

RESULT 11
US-09-135-166-5
; Sequence 5, Application US/09135166
; Patent No. 6083743
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,166
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-829 MTS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2950 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 334..2724
US-09-135-166-5

```

Alignment Scores:

```

Pred. No.: 2.03e-08 Length: 2950
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: Gaps: 44

US-09-914-168-2 (1-919) x US-09-135-166-5 (1-2950)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 322 AGGATACAATCGATGAAACAACTTCTAATCGCAAGTTTATTATTCGGTACGACACAGCT 381
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 382 GTGTTTGGCGCA-----CCTTTTGTGGCAAAAGATATTGCTGTGGATGCT 426
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 427 GTTCAAGGTGACTTAGAACACAAATCCGACCAAGTTTACCT----- 468
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 469 GTTCGTGCGCGTCAAGCTGCTGCTGCAATGATGTG---GCTAATATTGTCGCTCTTTA 525
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 526 TTCGTAGTGTGCGATTCGATGATGTAAGCGCATCAAGAGGCGGATGCTGTTTGT 585
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 586 ACGGTGTCGCTAAATCGATCATTTTCAGATGTTAAATCAAAAGGTAACCTGTTATTCCC 645
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 646 ACTGAGCAGCTTAAACAAACTTAGCTGCTAACGGGTTTAAAGTTGGCGATGTTTAA 702
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 703 -----ATTGCGAGAAAATTAATCAATTTGCCAAAAGTGTAAAGAGCAGCTATGCA 753
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 754 AGTGTAGTCGCTAT---AACGCAACAGTTGAACCTATTCTCAATAGCTACCAATAAT 810
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 811 CGCGCTGAAATTTTAAATCAATCAAGATCATAAACCAAAATTTGGCATCATTAAC 870
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 871 TTC-----AAGGGGAGCGAA---TCTGTTAGTAGCAGTACATTACAAGAA 912
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 913 CAATGGAATTACAACCTGATTCTTGGTGAAATTTATGGGAAATTAATTTGAAGTGCG 972
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 973 CAATTCGAGAAAGATTTCAGTCAATTCGT-----GATTAT 1008
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1009 TATTTAAATAAGGCTATGCCAAAGCACAATTTACTTAAACCGGATGTTTCAGCTAAATGAT 1068
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1069 GAAAAACAAAGTTAATGTACCACTTGTAAATGAAGGTTTACAGTATGACCTTCT 1128
QY 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1129 ACTGCACGCAATTAAGGTAATCTGGGAGGTATGCTCTGCGGAGCTTGAACCTTTACTTCA 1188

```



Qy 466 ThrValGluLeuThrAspGlyLeuLeuMetAspIleSerProIleGluPheSerAla 484  
Db 1189 GCATTACATTTAAATGATATTCGCGCGTAGTGATATTCGAGATGTAGAA 1239  
Qy 485 SerAsnLeuIleGlnAspLysLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1240 ---AATGCAATTAAGCAAACTTGGAGAACGCGTTACGGTAGCGCAACGGTAATAATCA 1296  
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 1297 GTACCTGATTTGATGCGCAATTAACATATAGCGATAACCTTGTGTGATCCTGGA 1356  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1357 CGAGCTTTAACTGTTGCGCAACTTCGCTTGAAGGAATACCGCTTCTGCTGATAGCACT 1416  
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
Db 1417 TTACGTCAGGAATCGGCCAACAGAGAACTTGGTATAATTCACAATTAGTTGAGTTA 1476  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1477 GGAATAATTCGCTTAGTCGACAGTTTCTTCGAAACAGTCGAAACCGAATTCATCCT 1536  
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
Db 1537 ATCAATGCTAGTAAATGAGTGGATGCTATATAAAGTCAAA----- 1581  
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592  
Db 1582 -----GAAACGTAACACGGGTAGTATCAACTTTGGTATTGGTTACGGT 1623  
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe-----GluHisAsnLeuIle----- 608  
Db 1624 ACAGAGAGTGTTAGTTATCAACCAAGTGTAAACCAAGATAATTTCTTGGACACAGG 1683  
Qy 609 -----AsnArgAspGlyTyrGlnGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1684 CGGCAGTAAAGTATAGCTGGTACGAAATAATGATTGTAGTACAGTGTCAATTTGGGTTAT 1743  
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1744 ACCGAGCCTATTTTACTAAAGATGGTGTAGTCTTGTGGAAATGTTTCTTTGAAAC 1803  
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSer 658  
Db 1804 TACGATAACTCTAAAGAGTACATCCCTCAACTATAAGCGTACGACTTACGGAAGTAAT 1863  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
Db 1864 GTTACTTTAGCTTTCCCTGTAATAAGAAATAACTCCTATTATGTAGGATTAGGTCATACC 1923  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr----- 686  
Db 1924 TATAATAAATTTAGTAACCTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 1974  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLysThrGlnAlaProProGluThrTrpGln 706  
Db 1975 CAATCAATGAAATTTAAAGGTAATGGCATTAAACAAAT----- 2013  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726  
Db 2014 -----GACTTTTGGATTTTCTTTTGGTTGG----- 2037  
Qy 727 AlaValHisLysThrValAlaAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2038 -----AATATAACAGCCTTAATAGAGGTATTTCCTCCAACT 2073  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2074 AAAGGGTTAAAGCAAGTCTTGGTGGAGGAGTTACTATTCCAGGTTCTGTATAACAAATAC 2133  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782

Db 2134 TACAAACTAAGTCGACATGTACAGGTTTCTTACCCATT----- 2172  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2173 GACAGAGATPCCTCTCGGTTGTATCTCAAAAGCATCTGCAGATATGCAAAATGGTTTT 2232  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820  
Db 2233 GGAACAACCGTTTACCGTTCTATCAAACTTATACAGCGGGTGCATCGTTTACATCGT 2292  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2293 GCTTTTGTCTTATGATGATGACCTTAACGCAATTTATGCCGAATATGCTAATGGTAGT 2352  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842  
Db 2353 GGTACTGGTACCTTTAAGAAAGATAGTTCTGTAT-----GTGATTGGTGGTAAATGCAATC 2406  
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856  
Db 2407 GCTACAGCTAGCGCAGAGTTAATTGTGCCCACTCCATTTGTGAGCGATTAAGACCCAAAT 2466  
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAlaAlaTyr----- 869  
Db 2467 ACGGTCGGAACCTCTTATTTCTGTATGCGCAAGTGTGGAAATCTAAATGGAATCA 2526  
Qy 870 -----AspLysGlyPheThrAsn 875  
Db 2527 GATAAAATGGATTAGAGCGCATGATTAAAGAAATGCTCCTGATTTATGCAAAATCAAGC 2586  
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArg 895  
Db 2587 CGTATTCGCGCCTCTACAGGTGTGCGATTCGAATGCAATCTCCTATTGGGCAATGGTA 2646  
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913  
Db 2647 TTCTCTTATGCCAACCAATTAATAAATATGAAATGATGATGTCGAACAGTTTCCAATTT 2706  
Qy 914 PheIleGlyThrProPhe 919  
Db 2707 AGTATTGGAGGTTCTTTC 2724  
RESULT 12  
US-08-942-046-5  
Sequence 5, Application US/08942046  
Patent No. 6264954  
GENERAL INFORMATION:  
APPLICANT: CHONG, Pele  
APPLICANT: THOMAS, Wayne  
APPLICANT: YANG, Yan Ping  
APPLICANT: LOOSMORE, Sheena  
APPLICANT: STA, Dwo Yuan Charles  
APPLICANT: KLEIN, Michel  
TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6TH Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1K7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/942,046  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2950 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 334..2724
; US-08-942-046-5

```

## Alignment Scores:

```

Pred. No.: 2,03e-08 Length: 2950
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 4 Gaps: 44

```

US-09-914-168-2 (1-919) x US-08-942-046-5 (1-2950)

```

QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 322 AGGATCAATCGATGAAACAACTCTAATCGCAAGTTTATTATTCGGTACGACACGACT 381
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 382 GTGTTGGCGCA-----CCTTTGTGGCAAAAGATATTCTGTTGGATGGT 426
QY 219 IleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 427 GTTCAAGGTGACTTAGAACCAAAATCCGACGACGTTTACCT----- 468
QY 239 LeuValAlaAlaAlaAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 469 GTTCGTCGCGGTACCGTGTGACTGACAAATGATGTG---GCTAATATTGTCCGCTCTTA 525
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 526 TTCGTAAGTGTGCGATTGCGATGATGCTGAAAGCGCATCAAGNAGCGATGCTGTGTGTT 585
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 586 AGCGTTGGTGCTAAATCGATCATTTTCAGATGTTTAAATCAAGTAAGTAACTCTGTTATCCC 645
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 646 ACTGAAGCAGCTTAAACAACTTAGATGCTAACGGTTTAAAGTTGGCGATGTTTA--- 702
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 703 -----ATTCCAGAAAATTAATGAATTTGCCAAAAGTGTAAGACGACCATATGCA 753
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 754 AGTGTAGTGCGCTAT---AACGCAACAGTTGAACCTATTGTCAATACGCTACCAATAAT 810
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 811 CGCGCTGAATTTAATCAATCAATGAAGATGATAAGCAAAATTTGGCATCATTA 870
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385

```

```

Db 871 TTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTACAAGAA 912
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 913 CAAATGAATTAACAACCTGATCTTGGTGGAAATTAATGGGAAATAAATTTTCAGAGTGGC 972
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 973 CAATTCGAGAAAGATTTCAGTCAATTCGT-----GATTAT 1008
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1009 TATTTAAATAATGGCTATGCCAACCAATTAACAAACGGATGTTTCAGCTAAATCAT 1068
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1069 GAAAAACAAAGTTAATGTAACCATTAAGCTTTTACAGTATGACCTCGT 1128
QY 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1129 AGTGCACGCATTATAGGTAATCTGGAGGTATCTCCGAGCTTGAACCTTTACTTTCA 1188
QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1189 GCATTACATTTAAATGATACTTCCGCCGTAGTATTCAGATATTCAGATCTAGAA----- 1239
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1240 ---AATGCAATTAAAGCAAACTTGGAGAACGCGGTTACGGTAGCGCAACGGTAAATTTCA 1296
QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1297 GTACCTGATTTTATGATGATGCAATAAACAATTAACGATAACCTCTGTTGTTGCTGCA 1356
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1357 CGACGTTTAACTGTCGCCCAACTTCCTTTGAAGGAATACCGCTTCTGCTGATGACACT 1416
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1417 TTACGTCAGAAATGCGCCCAACGAAGAACTTGGTATATAATTCACAATTAGTTGAGTTA 1476
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1477 GGAAAAATTCGCTTAGATGCTACAGGTTCTTCGAAACAGTCGAAACCCGAATTGATCCT 1536
QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1537 ATCAATGTTAGTAAATGATGAAGTGCATGCTGATATAATAAGTCAAA----- 1581
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1582 -----GAACGTAACACGGGTAGTATCAACTTTGTTGTTGTTGTTGTTGTTGTTGTT 1623
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1624 ACAGAGTGGTATTAGTTATCAAGCAAGTGTAAACAAGATAAATTTCTTGGCAACAGCG 1683
QY 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuLeuArgLeu 621
Db 1684 CGCGCAAGTATAGTGTGTACGAAAAATGATTATGTTGCTAGCTGCTCAATTTGGGTTAT 1743
QY 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1744 ACCGAGCCCTATTTTACTAAAGATGGTGAAGTCTTGGTGGAAATCTTTTCTTTGAAAC 1803
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis 658
Db 1804 TACGATAAATCTCTAAAGTACATCTCTTAATAAGCGTACGACTTACGGAAGTAAT 1863
QY 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671

```

```
Db 1864 GTTACTTTAGTGTTCCTCTGTAATAAATACTCTATTATGTAGGATTAGGTCATACC 1923
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr----- 686
Db 1924 TATAATAAAATAGTAACCTTTCCTCTAGAA-----TATAACCGTAATTTATATATT 1974
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706
Db 1975 CAATCAATGAAATTAAGGTAATAGGCAATTAACAACAAAT----- 2013
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2014 -----GACTTTTGATTTCTTCTTTGGTGG----- 2037
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2038 -----AACTATAACAGCCTTAATAGAGCTATTTCCTCAACT 2073
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2074 AAAGGGTTAAAGCAAGCTTGGTGGAGGAGTTACTATTCCAGGTTCTGTATAACAATAAC 2133
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2134 TACAACTAAGTCGATGTACAGGTTCTTACCATT----- 2172
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2173 GACAGAGATCACCTCTGGTGTGTATCTGCAAAAGCATCTGCAGGATATGCAATGGTTT 2232
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArg 820
Db 2233 GGAACAAGCGCTTACCTCTATCAAACTTATACAGCGGTGGCATCGTTTACGT 2292
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2293 GGTTCCTTATGTAGTAGTATTGGACCTAACCGAATTTATGCGGAATATGGTAAGTGTAGT 2352
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2353 GGTACTGTACTTTTAAAGAAGATAAGTCTGAT-----GTGATTGGTGGTAATGCAATC 2406
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2407 GCTACAGCTACGCGAGAGTTTAATTGTGCCAACTCCATTGTGAGCGATAAGAGCCAAAT 2466
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2467 ACGGTCCGAACCTCTTATTGTGTGATCGCGCAAGTGTGGTAATACTAAATGGAATCA 2526
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2527 GATAAAATGGATTAGAGCGATGTATTAAAGAGATTGCGCTGATTATGGCAAAATCAAGC 2586
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArg 895
Db 2587 CGTATTCGGCCTCTACAGGTGTCGGATTCCAATGGCAATCTCCTATTGGGCCATTGTA 2646
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2647 TTCTCTTATGCCAACCAATTAATAAATATGAAATGATGATGTCGAACAGCTTCCAATTT 2706
Qy 914 PheIleGlyThrProPhe 919
Db 2707 AGTATTGGAGGTTCTTTC 2724
```

## RESULT 13

US-08-433-522A-3

```
; Sequence 3, Application US/08433522A
; Patent No. 6013514
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
```

```
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,522A
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 374..2764
; US-08-433-522A-3
;
Alignment Scores: 2.07e-08 Length: 2984
Pred. No.: 172.00 Matches: 171
Score: 35.21% Conservative: 141
Percent Similarity: 19.30% Mismatches: 344
Best Local Similarity: 3.64% Indels: 230
Query Match: 3 Gaps: 44
DB:
;
US-09-914-168-2 (1-919) x US-08-433-522A-3 (1-2984)
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 362 AGGATACATCATGATAAATACTTCTAATCGCAAGTTTATTATTTCGTACGACACGACT 421
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 422 GTGTTTGGCGCA-----CCTTTTGGCAAAAGATATTCTGTGGATGGT 466
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 467 GTTCAAGGTGACTTAGAACACAAATCCGACGAAGTTTACCT----- 508
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 509 GTTCTGCGCGGTACCGCTGTGACTGACATGATGTG---GCTAATATTTCGCGCTTTTA 565
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 566 TTCGTAAAGTGTGTCGATTCGATGATGTAAGAGCGCATCAAGAGCGCATGTGCTTGTGTT 625
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla----Asp 288
Db 276 AspTyrArgAla-----ValGluValArgGlyGlyAla----Asp 288
```

Db 626 AGCGTTGTGGCTAAATCGATCATTTTCAGATGTTAAATCAAAAGGTAACCTCTGTTATTCC 685  
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308  
Db 686 ACTGAAGCCTTAAACAAACCTTACGATGCTAACCGGTTTAAAGTTGGCGATGTTTA--- 742  
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328  
Db 743 -----ATTGCGAGAAATAATTAATGAATTTGCGAAAGTGTAAGAGCAGCATATGCA 793  
Qy 329 PheAspGlyArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsn 345  
Db 794 AGTGAGTGGCTGATP---AACCCACAGTTGAACCTATTGCAATACGCTACCAATAAT 850  
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
Db 851 CGCGCTGAATTTAATTCAAATCAATGAAGATGATAAAGCAAAATTTGCGCATCACTA 910  
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
Db 911 TTC-----AAGGGGAACGAA---TCTGTGTAGTAGCAGTACATTACAGAA 952  
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
Db 953 CAATGGAATACACCTGATTCTTGGTGGAATTTATCGGGAATAAATTTGAAGGTGCG 1012  
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 1013 CAATTCGAGAAAGATTTCAGCTCAATTCGT-----GAUTAT 1048  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1049 TATTTAATATGCTATGCCAAAGCACAAATTTACTAAACGGATGTTTCAGCTAAATGAT 1108  
Qy 439 GlnValSerPheGluGlnSerSerArgThrGluProAlaGlnValAsp----- 456  
Db 1109 GAAAAACAAAGTAATGTAACCATGATGTAATGAAGGTTTACAGTAGTACCTTCGT 1168  
Qy 457 -----GluSerThrLeuGluProValIleGlu 465  
Db 1169 AGTCACGCATTATAGTAATCTGGGAGGTATGCTCCGAGCTTGAACCTTTACTTTCA 1228  
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
Db 1229 GCATTACATTTAAATGATACTTCCGCGTAGTCATATGCAGATGTAGAA----- 1279  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1280 ---AATGCAATTAAGCAAAACTTGGAGAACGCGGTTTACGGTAGCGCAACGGTAATTC 1336  
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 1337 GTACCTGATTTGATGATGCGCAACTTCGCTTTGAAGGAATACCGTTCTGCTGATAGCACT 1456  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1397 CGAGGTTTAACTGTTGCCCAACTTCGCTTTGAAGGAATACCGTTCTGCTGATAGCACT 1550  
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
Db 1457 TTACGTGAGGAATGCGCCCAACGAAGGAAGAACTTGTGTATTAATTCACAATTTAGTTGAGTTA 1516  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1517 GGAATAATTCGCTTAGATCGTACAGGTTTCTCGAACAACGCTGAAACCGAATTTGATCCT 1576  
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
Db 1577 ATCAATGGTAGTAATGATGAAGTGGATGTCGTATATAAAGTCAA----- 1621  
Qy 575 ValPheValAlaSerAspLysProArgAspGlyClnIle-----GlyLeuGlyTyrGly 592  
Db 1622 -----GAACGTAAACGGGTAGTATCAACTTTTGGTATTGGTTACGGT 1663

Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608  
Db 1664 ACAGAGAGTGGTATTAGTATTCAAGCAAGCTGTAAACAAAGATAAATTTCTGGGAACAGGG 1723  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1724 GCGGCGAGTAAGTATAGCTGGTACGAAATGATTTAGTACGAGTGTCAATTTGGTTAT 1783  
Qy 622 SerGlu-----AspLysLysValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1784 ACCGAGCCCTATTTTACTAAAGATGGTGAAGTCTGTGGAAATGTTTCTTCTTCAAAAC 1843  
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
Db 1844 TACGATACTCTAAAGTGATACATCCCTCACTAATAAGCGTACGAGTACCGAAGTAAT 1903  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
Db 1904 GTTACTTTAGGTTTCCCTGTAATGAAATCAACTCCTATTATGTAGGATTAGGTCAATACC 1963  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrPAsnArgThr--- 686  
Db 1964 TATATAAATAAGTAACTTTTCTCTAGAA-----TATAACCGTAATTTATATATT 2014  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysProSerGlnAlaProProGluThrTrpGln 706  
Db 2015 CAATCAATGAAATTTAAAGGTAATGGCAATTAACAAAT----- 2053  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726  
Db 2054 -----GACTTTGATTTTTTCTTTTGGTGG----- 2077  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2078 -----AACTATAACAGCCTTAATAGAGCTATTTCCCAACT 2113  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2114 AAAGGGTTAAAGCAAGCTTGTGGTGGAGCTTACTATTCCAGGTTCTGATAACAATAC 2173  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 2174 TACAACTAAGTCAGATGATACAGGTTTCTACCCATTA----- 2212  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2213 GACAGAGATCACCTCTCGGTTGTATCTGCAAAAGCATCTCGAGGATATGCAATGGTTTT 2272  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyAspGlnSerIleArg 820  
Db 2273 GGAACAACAGCGTTTACCGTTCTATCAAACTTATACAGCGGTGCGCATCGTTCAATACGT 2332  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2333 GGTTCCTTATGCTAGTATGAGCTTACGCAATTTATGCCGAATATGCTAAATGGTAGT 2392  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842  
Db 2393 GGTACTGGTACTTTTAAAGAGATAAGTCTCTGAT-----GTGATTGCTGTAATGCAATC 2446  
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp 856  
Db 2447 GCTACAGCTAGCGCAGAGTTAATTTGTGCCAACTCCATTTGTGAGCGCATAGAGCCAAAT 2506  
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869  
Db 2507 ACGTCCGAAACCTCTTATTTGTTGATGCGGCAAGTGTTTTGGAAATACTAAATGCAATCA 2566  
Qy 870 -----AspLysGlyPheThrAsn 875  
Db 2567 GATAAAATGGATTACAGAGCGGATGTTATTAANAAGATTGCCTGATTATGCAATCAAGC 2626

Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrPalaSerProValGlyGlnValArg 895  
Db 2627 CGTATTCGCGCTACAGGTGTCGATTCCCAATGCAATCTCCATTGGGCAATTGGTA 2686  
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913  
Db 2687 TTCCTTATGCCAAACCAATTAATAAATATGAAATGATGATGTCGCAACAGTTCCAATTT 2746

Qy 914 PheIleGlyThrProPhe 919

Db 2747 AGTATTGGAGGTCTCTTC 2764

## RESULT 14

US-09-135-166-3

; Sequence 3, Application US/09135166

; Patent No. 6083743

; GENERAL INFORMATION:

; APPLICANT: CHONG, Pele

; APPLICANT: THOMAS, Wayne

; APPLICANT: YANG, Yan Ping

; APPLICANT: LOOSMORE, Sheena

; APPLICANT: SIA, Dwo Yuan Charles

; APPLICANT: KLEIN, Michel

; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN

; NUMBER OF SEQUENCES: 55

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6TH Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1K7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/09/135,166

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/433,522

; FILING DATE: 12-SEP-1995

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2984 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 374..2764

US-09-135-166-3

## Alignment Scores:

Pred. No.:	2,07e-08	Length:	2984
Score:	172.00	Matches:	171
Percent Similarity:	35.21%	Conservative:	141
Best Local Similarity:	19.30%	Mismatches:	344
Query Match:	3.64%	Indels:	230
DB:	3	Gaps:	44

US-09-914-168-2 (1-919) x US-09-135-166-3 (1-2984)

Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198  
Db 362 AGGATACAACTCGATGAAAAAACTTCTAATCGCAAGTTTATTATTCGGTAGCAACGACT 421  
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218  
Db 422 GTGTTTGGCGCA-----CCTTTTGGCAAAAGATATTTCGTGGGATGGT 466  
Qy 219 IleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238  
Db 467 GTTCAAGGTGACTTAGAACACAAATCCGAGCAAGTTTACCT----- 508  
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleLeuArgAsnSer 258  
Db 509 GTTCGTGCGCGGTGAGCTGCTGACTGACAAATGATGATG--GCTAATATTCGCGCTCTTTA 565  
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275  
Db 566 TTCGTAAAGTGGTTCGATTTCGATGATGTAAGCGCATCAAGAGCGGATGTGCTTTGTT 625  
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288  
Db 626 ACCGTTGTCGCTAAATCGATCATTTTCAGATGTTTAAATCAAGGTAACTCTGTATTCCC 685  
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308  
Db 686 ACTGAAGCACTTAAACAAAACTTAGATGCTAACGGGTTTAAAGTTGCGGATGTTTA-- 742  
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328  
Db 743 -----ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGAAGAGGACTATGCA 793  
Qy 329 PheAspGlyArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsn 345  
Db 794 AGTGTAGGTGCTAT---AACGCAACAGTTGAACCTTGTCTCACTACGCTACCAATAAT 850  
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
Db 851 CGCGCTGAAATTTAATTCAAATCAATGAAGATGAATAAGCAAAATTTGGCATCATTAAC 910  
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrValAsnMetGlyGluAlaTyr----- 402  
Db 911 TTC-----AAGGGAACGAA---TCTGTTAGTAGCAGCTACATTACAAGAA 952  
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
Db 953 CAAATCGAATTACAACTGATTTCTTGGTGCAAAATTTATGGGAAATAAATTTGAAGTCCG 1012  
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 1013 CAATTCGAGAAAGATTTGCGAGTCAATTCGT-----GATTAT 1048  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1049 TATTTAAATAATGGCTATGCCAACGACAAATTTACTAAAACGGATGTTCAGCTAAATGAT 1108  
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456  
Db 1109 GAAAAACAAAAGTTAATGTAAACCATGTGTAATGAAGGTTTACAGTATGACCTTCGT 1168  
Qy 457 -----GluSerThrLeuGluProValIleGlu 465  
Db 1169 AGTGCACGCAATTATAGTAATCTGGGAGGTATGCTGCCGCGAGCTTGAACCTTTACTTTCA 1228  
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
Db 1229 GCATTACATTTAAATGATCTTTCCCGCGTAGTATGATATTCAGATGTAGAG----- 1279  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1280 ---AATGCAATTAAAGCAAAACTTGGAGAACGCGGTTACCGTAGCGCAACGGTAATTC 1336

```
QY 505 MetProAsp-----AspArgValLeuAlaTleAsnHisAspAspGlyValAsn 520
Db 1337 GTACCTGATTGTGATGCAAAATAAAACATTAGCGATAACCCCTGTTGTTGATGCTGGA 1396
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1397 CGAGCTTTAACTGTTCCCAACTTCGTTTGAAGGAATAACCCCTTCTGCTGATAGCACT 1456
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1457 TTACGTCAGGAAATGCGCCACACAGAAGAACTTGCTATATAATTCACAAATAGTTGAGTTA 1516
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1517 GGAATAATTCGCTTAGCTAGCTACAGCTTCTTCGAAACAGTCGAAACCGAAATGATCCT 1576
QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1577 ATCAATGGTAGTAATGATGAAGTGGATGCTGATATAAAGTCAAA----- 1621
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1622 -----GAAAGTAAACACGGGTAGTATCAACTTTGGTATTGTTACGGT 1663
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe-----GluHisAsnLeuIle----- 608
Db 1664 ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTAAACACAGATAAATCTCTGGGAACAGGG 1723
QY 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1724 GCGGAGTAGTATAGCTGGTACGAAATAATGATGTTACGAGTGTCAATTTGGGTTAT 1783
QY 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1784 ACCGAGCCCTATTTACTAAAGTGGTGAAGTCTGTGGTGAATGTTTCTTTGAAAC 1843
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGluValPheGlyHisSer 658
Db 1844 TACGATAACTCTAAAGTGATACATCCCTAACTAATAGCTAGCAGTACCGAAGTAAT 1903
QY 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1904 GTTACTTTAGGTTCCCTGTTAAAGTAAATAACTCTTATTAATGATGAGGATTAGTCAACC 1963
QY 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrPsnArgThr----- 686
Db 1964 TATAATAAAATTAAGTAACTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 2014
QY 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 2015 CAATCAATGAATTTAAAGGTAATGGCATTAACAAACAAAT----- 2053
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2054 -----GACTTTGATTTTCTTTTGGTTGG----- 2077
QY 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2078 -----AACTATAACAGCCTTAATAGAGGCTATTTCCCAACT 2113
QY 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2114 AAAGGGTTAAAGCAAGCTTTGGTGGACGAGTTACTATTCACGGTCTTGATAACAAATAC 2173
QY 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2174 TACAAACTAAGTCAGATGCTACAGGTTTCTACCCATTA----- 2212
QY 783 AsnArgAlaHis-----GlnMetThrGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2213 GACAGAGATCACCTCTGGGTTGATCTGCAAAAGCATCTGAGGATATGCAATGGTTTTT 2272
QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
```

```
Db 2273 GGAACAACAGCGTTTACCGTTCTCTATCAACAGCGGTGGCATCGGTTTCATTACGT 2332
QY 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2333 GGTTTTGCTTATGCTAGTATTGGACCTAACCGAAATTTATGCCGAATATGGTAATGGTAGT 2392
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2393 GGTACTGGTACTTTTAAAGAAGATAAGTTCTGTAT-----GTGATTGGTGGTAATCAATC 2446
QY 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2447 GGTACAGCTAGCGCAGAGTTAATTTGTGCCAATCCATCTCTATTGGGCCATTGGTA 2686
QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2507 ACGGTCCGAAGCTCTCTATTGTTGATCGGCAAGTGTTCGAATACTAAATGGAATCA 2566
QY 870 -----AspLysGlyPheThrAsn 875
Db 2567 GATAAAATGATTAGACAGCGATGTTATTAAGAGATTGCCCTGATTATGGCAATCAAGC 2626
QY 876 AspThrLysIleGlyAlaGlyValArgThrAlaSerProValGlyGlnValArg 895
Db 2627 CGTATTCGCGCTCTACAGGTGTCGGATTCCAATGGCAATCTCCTATTGGGCCATTGGTA 2686
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuIlePhe 913
Db 2687 TCTCTCTATGTCACCAACCAATTAATAAATATGAAATGATGATGTCGAACAGTTTCCAATTT 2746
QY 914 PheIleGlyThrProPhe 919
Db 2747 AGTATTGAGGTTCTTTC 2764
```

```
RESULT 15
US-08-942-046-3
; Sequence 3, Application US/08942046
; Patent No. 6264954
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/942,046
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-Sep-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb
; TELECOMMUNICATION INFORMATION:
```

Db	1013	CAATTCGAGAAAGATTTCGACGTCAAATTCGT-----GATTAT	1048
Qy	419	TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp	438
Db	1049	TATTTAAATTAATGGCTATGCCAAAGCACAAATTACTAAACCGGATGTTTCAGCTAAATGAT	1108
Qy	439	GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp-----	456
Db	1109	GAATAACAAAGCTTAATGTAACCATTTGATGTAATGAGGTTTACAGATGACCTTCGT	1168
Qy	457	-----GluSerThrLeuGluProValIleGlu	465
Db	1169	AGTCACGCATTAATAGTAACTCTGGGAGTATGTCTGCCGACCTTGAACCTTTACTTCA	1228
Qy	466	ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla	484
Db	1229	GCATTACATTTAAATGATACTTTCGCCCGTAGTATATTGCAGATGTAGAA-----	1279
Qy	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp	504
Db	1280	---AATGCAATTAAGACAAACTTCGAGAACCGGGTACGGTAGGCCAACGGTAATAATCA	1336
Qy	505	MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn	520
Db	1337	GTACCTGATTTTGATGATGCAATAAATACATTAGCGATAACCCCTGTGTTGATGCTCGA	1396
Qy	521	ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla	533
Db	1397	CGAGCTTTAACTGTTGCCCAACTTCGCTTTGAAGGAATAACCGTTTCTGCTGATAGCACT	1456
Qy	534	ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu	550
Db	1457	TTAGCTCAGGAATGCGCCAAACAAGAACTTGGTATAATTCACAATTAGTTGAGTTA	1516
Qy	551	Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro	562
Db	1517	GGAAAAATTCGCTTAGATCGTACAGAGTTCTTCGAAACACGTGCAAAACCGAATTGATCCT	1576
Qy	563	AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr	574
Db	1577	ATCAATGCTAGTAATGATGAAGTGGATGCTGATATATAAAGTCAAA-----	1621
Qy	575	ValPheValAlaSerAspLysProArgspGlyGlnIle-----GlyLeuGlyTrpGly	592
Db	1622	-----GAACGTAAACACGGGTAGTATCAACTTTGGTATTGGTTACGGT	1663
Qy	593	SerAspThrGlyThrArgLeuValThrLysPhe--GluHisAsnLeuIle-----	608
Db	1664	ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTAAACAAATAGATATTTCTTGGGAACAGGG	1723
Qy	609	-----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu	621
Db	1724	CGGCGACGTAAATAGTGTGTACGAAATAATGATTATGGTACCGAGTCAATTTGGGTTAT	1783
Qy	622	SerGlu-----AspLysGlyValLysLeuTyrAlaThrLysProLeuSerHis	638
Db	1784	ACCGAGCCCTATTTTACTAAAGATGCTGTAAAGTCTGGTGGGAAATGTTTTCTTTGGAAAC	1843
Qy	639	ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheIleHisSer	658
Db	1844	TACGATAACTCTAAAGTGATACATCCTCTAACTATTAAGCGTACGACTTACGGAAGTAAT	1903
Qy	659	---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis---	671
Db	1904	GTTACTTTAGTTTTCCTCTGTAATGAAAAATAACTCCTATTATGTAGGATATAGGTCATACC	1963
Qy	672	----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr-----	686
Db	1964	TATAATAAAATTAGTAACCTTTCCTAGAA-----TATAACCGTAATTTATATATT	2014
Qy	687	TyrSerLeuArgThrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln	706

```
Db 2015 CAATCAATGAATTAAGGTAATGGCATTAAAAACAAAT----- 2053
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
    |||||
Db 2054 -----GACITTTGATTTTCTTTGGTTGG----- 2077
QY 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
    |||
Db 2078 -----AACTATAACAGCCTTAATAGAGGCTATTTCCCAACT 2113
QY 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
    |||||
Db 2114 AAGGGGTTAAAGCAAGCTCTGGTGGAGGAGTACTATTCAGGTTCTGTATAACAAATAC 2173
QY 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
    |||||
Db 2174 TACAACTAAGTCAGATGATACAGGGTTTCTACCCATTA----- 2212
QY 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
    |||||
Db 2213 GACAGAGATCACCTCTGGTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGGTTT 2272
QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
    |||||
Db 2273 GGAACAAGCGTTTACCGTCTCTATCAACTATACACGGGTGCATCGGTTCAATTACGT 2332
QY 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
    |||||
Db 2333 GGTTCCTTATGCTAGTATTGAGCAATTAACGCAATTTATGCGCAATATGGTAATGGTAGT 2392
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
    |||||
Db 2393 GGTACTGGTACTTTTAAGAAGATAAGTCTCTGAT-----GTGATTGGTGTAAATGCAATC 2446
QY 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
    |||||
Db 2447 GCTACAGCTAGCGCAGAGTAATTTGTGCCAACTCCATTTGTGACGATAGAGCCAAAT 2506
QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
    |||||
Db 2507 ACGTCGGAACCTCTTATTATTGTTGATGCGGCAAGTGTTTGGAATACTAAATGGAATCA 2566
QY 870 -----AspLysGlyPheThrAsn 875
    |||||
Db 2567 GATAAATGATAGAGCGGATGTATTAAGAGATTGCTGTGATTGGCAAAATCAAGC 2626
QY 876 AspThrLysIleGlyAlaGlyValArgTyrAlaSerProValGlyGlnValArg 895
    |||||
Db 2627 CGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCTCTCTATTTGGCCATTGGTA 2686
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
    |||||
Db 2687 TTCTCTTATGCCAAACCAATTAATAAATATGAAAAATATGATGATGTCGAACAGTTCCAATT 2746
QY 914 PheIleGlyThrProPhe 919
    |||||
Db 2747 AGTATTGGAGGTTCTTTC 2764

RESULT 16
US-08-433-522A-55
; Sequence 55, Application US/08433522A
; Patent No. 6013514
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuen Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
```

```
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,522A
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2987 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-433-522A-55

Alignment Scores:
Pred. No.: 2,08e-08 Length: 2987
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 3 Gaps: 44

US-09-914-168-2 (1-919) x US-08-433-522A-55 (1-2987)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
    |||||
Db 377 AGGATACATCGATCAAAAACCTCTAATCGCAAGTTTATTATTTCGGTACGACACGACT 436
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
    |||||
Db 437 GTGTTTGGCGCA-----CCTTTTGTGCAAAAGATATTTCGTGGATCGT 481
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
    |||||
Db 482 GTTCAAGGTGACTTAGAACACAAATCCGACGACAGTTTACCT----- 523
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
    |||||
Db 524 GTTCGTGCGCGTACAGCGTGTGCTGACTGACAATGATGTG---GCTAATATTGTCGCTCTTTA 580
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
    |||||
Db 581 TTCGTAAGTGTGCGATTTCGATGATGTAAGAGCGCATCAAGAGCGCATGCTGCTTGTGTT 640
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
    |||||
Db 641 AGCGTTGTGGCTAAATCGATCATTTTCAGATGTTAAATCAAGAGGTAACCTGTTATTCCC 700
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
    |||||
Db 701 ACTGAAGCAGCTTAAACAAACTTAGATCTAACGGGTTTAAAGTTGGCGATGTTTAA--- 757
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
    |||||
Db 758 -----ATTCGAGAAAAAATAAATGAATTTGCAAAAGTGTAAAAAGACACATGCA 808
QY 329 PheAspGlyArgTyrProLeuAspArgSerValAspValIle-----LeuProAspAsn 345
    |||||
```



Db 809 AGTGAGTTCGCTAT---AACGCACACAGTTGAACCTATTGCTAAATACGCTACCAATAAT 865  
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
Db 866 CGCGCTGAATTTAAATCAATCAATCAAGATGATAAGCAAAATTCGCATCACTACT 925  
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
Db 926 TTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTACAAGAA 967  
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
Db 968 CAAATGGAAATTACACACCTGATCTTGGTGGAAATTTATGGGAAATAAATTTGAAGGTGCG 1027  
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 1028 CAATTCGAGAAAGATTTCAGCTCAATTCGT-----GATTAT 1063  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
Db 1064 TATTTAAATAATGGCTATGCCAAAGCACAAATTTACTAAACGGATGTTTCAGCTAAATGAT 1123  
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456  
Db 1124 GAAAAACAAAAGTTAATGTAAACCATTCATGATGTAATGAAGGTTTACAGTATGACCTTCGT 1183  
Qy 457 -----GluSerThrLeuGluProValIleGlu 465  
Db 1184 AGTGCACCATATAGGTAATCTGGGAGGTATGTCGCGGAGCTTGAACCTTACTTTCA 1243  
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGlnPheSerAla 484  
Db 1244 GCATTACATTTAAATGATGACTTTCGCGCGTAGTGATATTCAGAGATGAGAA----- 1294  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1295 ---AATGCAATTTAAAGCAAACTTGGAGACGCGTTTACGGTAGCGCAACGGTAAATTC 1351  
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 1352 GTACCTGATTTGATGATGCAATAAACAATTAGCGATAACCCCTGTGTTGATGCTGGA 1411  
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
Db 1412 CGAGCTTTAACTGTTGCGCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTGATAGCCT 1471  
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
Db 1472 TTACGTCAGGAATGCGGCCAACAGAGAACTGGTATATTCATAATTCATAATTTAGTTAGGTTA 1531  
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
Db 1532 GGAATAATTCGTTAGATCGTACAGTTTCTTCGAAACAGTCCGAAACCGAATTCATCCT 1591  
Qy 563 AlaAsp-----ValTyrGlnSerLysValProLeuTyr 574  
Db 1592 ATCAATGGTAGTAATGAAAGTGGATGCTGTATATAAAGTCAAA----- 1636  
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592  
Db 1637 -----GAACGTAAACAGGGTAGTATCAACTTTGGTATTGGTTACCGT 1678  
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608  
Db 1679 ACAGAGAGTGGTATTAGTTATCAACCAAGTGTTAAACAAGATAAATTTCTTGGGAACAGG 1738  
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
Db 1739 GCGGCAGTAAGTATAGCTGGTACGAAAAATGATTGTGTACGAGTCTCAATTTGGGTTAT 1798  
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
Db 1799 ACCGAGCCCTATTATTACTAAAGATGGTGTAGCTCTTGGTGGAAATGTTTCTTTGAAAC 1858

Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
Db 1859 TACGATAACTCTAAAGTCATACCTCTAATACTATAAGCGTACGACTTACGGAAGTAAT 1918  
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
Db 1919 GTTACTTTTAGGTTCCCTGTAATGAAATAACTCCTATTATGATAGGATTAGCTCATACC 1978  
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr----- 686  
Db 1979 TATAATAAAATTTAGTAACCTTTCCTCTAGAA-----TATAACCGTAATTTATATATT 2029  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLysThrGlnAlaProProGluThrTrpGln 706  
Db 2030 CAATCAATCAATTTAAAGGTAATGCGATTAAACAAT----- 2068  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726  
Db 2069 -----GACTTTTGATTTTCTTTTGGTTGG----- 2092  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 2093 -----AACTATAACAGCCTTAATAGAGGCTATTTCCTCCAACT 2128  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
Db 2129 AAAGGGGTTAAAGCAAGTCTTGGTGCAGGATTACTATTCCAGGTTCTGATAACAATAAC 2188  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 2189 TACAACAACTAAAGTCAGATGCTACAGGTTTCTACCCATTA----- 2227  
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800  
Db 2228 GACAGAGATCAGCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCCAATGGTTTT 2287  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyGlyGlnSerIleArg 820  
Db 2288 GGAACAACAGCGTTTACCGTTCTCTCAAACTTATACAGCGGGTGCATCGTTCACTACCT 2347  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2348 GCTTTTGTCTTATGGTAGTATTGGACCTAACGCAATTTATGCCGAATATGTTATGGTAGT 2407  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842  
Db 2408 GGTACTGGTACTTTTAAAGNAGATAAGTTCTGAT-----GTGATGGTGGTAAATGCCAATC 2461  
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856  
Db 2462 GCTACAGCTAGCGCAGAGTTAATTTGCCCAACTCCATTTGTGAGCGATAAGAGCCAAAAT 2521  
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869  
Db 2522 ACGGTCGCAACCTCTCTATTGTTGATGCGCAAGTGTGGAATACTAAATGGAAATCA 2581  
Qy 870 -----AspLysGlyPheThrAsn 875  
Db 2582 GATAAAATGGATTAGAGAGCGATGATTAAANAAGATTGCTGATTTATGGCAATCAAGC 2641  
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArg 895  
Db 2642 CGTATTTCGCGCTCTACAGGTCTCGGATTCCAATGSCAATCTCCTATTGGGCCATTGGTA 2701  
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913  
Db 2702 TTCTCTTATGCCAAACCAATTAATAAATATGAAAAATGATGATGTCGAACAGTTTCCAATTT 2761  
Qy 914 PheIleGlyThrProPhe 919  
Db 2762 AGTATTGGAGGTTCTTTC 2779

## RESULT 17

US-09-135-166-55  
 ; Sequence 55, Application US/09135166  
 ; Patent No. 6083743  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHONG, Pele  
 ; APPLICANT: THOMAS, Wayne  
 ; APPLICANT: YANG, Yan Ping  
 ; APPLICANT: LOOMORE, Sheena  
 ; APPLICANT: SIA, Dwo Yuan Charles  
 ; APPLICANT: KLEIN, Michel  
 ; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN  
 ; NUMBER OF SEQUENCES: 55  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sim & McBurney  
 ; STREET: 6TH Floor, 330 University Avenue  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5G 1R7  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/135,166  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/433,522  
 ; FILING DATE: 12-SEP-1995  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: STEWART, Michael I  
 ; REGISTRATION NUMBER: 24,973  
 ; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (416) 595-1155  
 ; TELEFAX: (416) 595-1163  
 ; INFORMATION FOR SEQ ID NO: 55:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2987 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear

US-09-135-166-55

## Alignment Scores:

Pred. No.:	2,08e-08	Length:	2987
Score:	172.00	Matches:	171
Percent Similarity:	35.21%	Conservative:	141
Best Local Similarity:	19.30%	Mismatches:	344
Query Match:	3.64%	Indels:	230
DB:	3	Gaps:	44

US-09-914-168-2 (1-919) x US-09-135-166-55 (1-2987)

QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerGlnSerGlyGluThrSerAla 198  
 Db 377 AGGATACAATCGATGAAAAAATCTTAATCGCAAGTTTATTATTCGGTAGCAGCAACT 436  
 QY 199 IleGlySerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLysGluAsp 218  
 Db 437 GTGTTGCCCA-----CCTTTTGGCAAAAGATATTCGTGTGGATGCT 481  
 QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238  
 Db 482 GTTCAAGGTGACTTAGAACCAAAATCCGACCAAGTTTACCT----- 523  
 QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleLeuArgAsnSer 258  
 Db 524 GTTCGTGCCGTGAGCGGTGACTGACAAATGATGTG--GCTAAATATTGTCGCTCTTTA 580

QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275  
 Db 581 TTCGTAAGTGGTTCGATTCGATGATGTAAGCGCATCAAGAGCGATGCTGTTGTTGT 640  
 QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288  
 Db 641 ACGGTTGCGCTAAATCGATCATTTTCAGATGTTAAATCAAAGGTAACTGTTATTCC 700  
 QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308  
 Db 701 ACTGAAGCACTTAAACAAACTTAGATCTAACGGGTTTAAAGTTGGGATGTTTAA--- 757  
 QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328  
 Db 758 -----ATTCGAGAAAAATTAATCAATTTGCCAAAAGTGTAAAGAGCAGCATATG 808  
 QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345  
 Db 809 AGTGTAGTCGCTAT---AACGCAACAGTTGAACCTATTGTCAATACGCTACCAAAATA 865  
 QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
 Db 866 CGCCTCAAAATTTTAAATCAATCAATCAAGATGATAAAGCAAAATTTGGCATCATTA 925  
 QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
 Db 926 TTC-----AAGGGGACGAA---TCTGTTAGTAGCAGTACATTTACAAGAA 967  
 QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
 Db 968 CAAATGGAATTACAACCTGATTTCTGGTGCAAAATTTATGGGAAATTAATTTGAAGTGC 1027  
 QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
 Db 1028 CAATTCGAGAAAGATTTCAGTCAATTCGT-----GATTAT 1063  
 QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
 Db 1064 TATTTAAATTAATGGCTATGCCAAAGCACAAATTTACTAAAACGGATGTTCCAGCTAA 1123  
 QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456  
 Db 1124 GAAAAACAAAGTTAATGTAACCATTTGATGTAATGAAGGTTTACAGTATGACCTTCGT 1183  
 QY 457 -----GluSerThrLeuGluProValIleGlu 465  
 Db 1184 AGTGCACGCATTATAGGTAATCTGGGAGGTATGTCGCGAGCTTCAACCTTTACTTTCA 1243  
 QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
 Db 1244 GCATTACATTTAAATCATACTTTCCGCCCTAGTATGATTTGCAGATGTAGAA----- 1294  
 QY 485 SerAsnLeuIleGlnAspLysLysLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
 Db 1295 ---AATGCAATTAAGCAAAACTTCGAGAACCGGTTACGGTAGCGCAACGTAATTTCA 1351  
 QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
 Db 1352 GTACCTGATTTTGATGATGCAAAATAAAACATTAGCGATACCCCTTGTGTTGATCGTGA 1411  
 QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
 Db 1412 CGAGCTTTAACTGTCGCCAACTTCGCTTTGAAGCAATAACCGTTTCTGCTGATAGCACT 1471  
 QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
 Db 1472 TTAGCTCAGGAAATCGGCCAACAAAGAACTTGGTATATAATTCACAAATTAGTTGAGTTA 1531  
 QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
 Db 1532 GAAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAAACAGTCGAAAAACCGAATTTGATCCT 1591

```
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
   :::: |||||::: |||
Db 1592 ATCAATGGTAGTAATGATGAAGTGGATCGTATATAAAGTCAAA----- 1636
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
   :::: ||| ||| |||||::: |||
Db 1637 -----GAACGTAAACACGGGTAGTATCAACTTTGGTATTGGTTACGGT 1678
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
   :::: |||||::: ||| |||
Db 1679 ACAGAGTGGTATTAGTATCAAGCAAGTGTAAACAAAGATAATTTCTTGGGACACAGG 1738
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
   :::: |||||::: |||
Db 1739 GCGGCAGTAAGTATAGTGGTACGAAATAATGATTGTAGGAGTGTCAATTTGGGTTAT 1798
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
   :::: |||||::: ||| |||||
Db 1799 ACCGAGCCCTATTTTACTAAAGATGGTGAAGTCTTGGTGGAAATGTTTCTTTGAAAC 1858
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSer 658
   ||| :::: ||| :::: |||||::: |||
Db 1859 TACGATRACTAAAGGTATACATCTCTAACTATATAGCGTACGACTTACGGAAGTAAT 1918
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
   ||| :::: |||||::: |||
Db 1919 GTTACTTTAGGTTTCCCTGTAATAAGAAATAACTCTATTATGTAGGATTAGGTCATACC 1978
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686
   :::: |||||::: |||||
Db 1979 TATAATAAATTAGTAAGTCTTGCTCTAGAA-----TATACCCGTAATTTATATATT 2029
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
   |||||::: |||||::: |||
Db 2030 CAATCAATGAATTTAAGGTAATGGCATTAACAAAT----- 2068
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuLeuAlaGlyVal 726
   :::: |||||::: |||
Db 2069 -----GACTTTTGATTTTCTTTTGGTGG----- 2092
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
   ||| :::: ||| :::: |||||
Db 2093 -----AACTATAACAGCCTTATAGAGGCTATTTCCTCCAACT 2128
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
   :::: |||||::: |||
Db 2129 AAAGGGTTAAAGCAAGCTTTGGTGGACGAGTTACTATTCCAGGTTCTGATAACAAATAC 2188
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
   ||| :::: ||| |||
Db 2189 TACAACTAAGTCAGATGTACAGGGTTTCTACCCATTA----- 2227
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
   :::: ||||| ||| :::: |||||
Db 2228 GACAGAGATCACCTCTGGTGTATCTGCAAAAGCATCTGCAGGATATGCAATGGTTTTT 2287
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyAspGlnSerIleArg 820
   :::: |||||::: ||| |||||
Db 2288 GGAACAACAGCGTTTACCCTTCTATCAAACTTATACAGCGGGTGGCATCGGTTTCATTACGT 2347
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
   :::: |||||::: |||
Db 2348 GGTTCCTTATGTAGTATTGGACCTAACGCAATTTATGCCAATATGGTAATGGTAGT 2407
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
   ||||| ||||| :::: |||||
Db 2408 GGTACTGTACTTTTAAAGAGATAGTTCTGAT-----GTGATGGTGGTAATCCAATC 2461
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
   ||| :::: |||||
Db 2462 GCTACAGCTAGCGCAGAGTTAATTTGTGCGCAACTCCATTTGTGAGGATAGAGCCAAAT 2521
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
```

```
Db 2522 ACGGTCCGAACCTCTCTATTATTGTTGATCGGCAAGTGTTTGGAATPACTAAATGGAATCA 2581
   :::: ||| :::: ||| ||| ::::
Qy 870 -----AspLysGlyPheThrAsn 875
   ||| ||| ::::
Db 2582 GATAAAATAGATTAGAGAGCGATGATTAAAGAGATTGCCTGATTATGGCAATCAAGC 2641
   :::: |||||::: ||| |||||
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArg 895
   :::: |||||::: ||| |||||
Db 2642 CGTATTCGCGCTCTACAGGTGTCGGATTCCAAATGGCAATCTCTATTGGGCCATTGGTA 2701
   :::: |||||::: ||| |||||
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
   ||| :::: ||| ||| ::::
Db 2702 TTCTCTTATGCAACCAATTAATAAATATGAAATATGATGTGCAACAGTTCCAATTT 2761
   :::: ||| ||| ::::
Qy 914 PheIleGlyThrProPhe 919
   ||||| |||
Db 2762 AGTATGGAGGTTCTTTC 2779

RESULT 18
US-08-942-046-55
: Sequence 55, Application US/08942046
: Patent No. 6264954
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: THOMAS, Wayne
: APPLICANT: YANG, Yan Ping
: APPLICANT: LOOSMORE, Sheena
: APPLICANT: SIA, Dwo Yuan Charles
: APPLICANT: KLEIN, Michel
: TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 6TH Floor, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/942.046
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/433,522
: FILING DATE: 12-SEP-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 55:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2987 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-942-046-55

Alignment Scores:
Pred. No.: 2.08e-08 Length: 2987
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
```

DB: 4 Gaps: 44  
US-09-914-168-2 (1-919) x US-08-942-046-55 (1-2987)  
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlnSerGlyGluThrSerAla 198  
DB 377 AGGATACAACTCGATGAAAAAATCTTCTAATCGCAAGCTTTATTATTTCGGTACGACACGACT 436  
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218  
DB 437 GTGTTGCGCGCA-----CCTTTGTGGCAAAAGATATTTCGTGGATCGT 481  
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238  
DB 482 GTTCAAGGTGACTTAGAACAAACAAATCGACGACAGTTTACCT----- 523  
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258  
DB 524 GTTCGTGCGGTGACGCTGACTGACCAATGATG---GCTAATATTGTCCGCTCTTTA 580  
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275  
DB 581 TTCGTAAAGTGGTCGATTCGATGATGTAAGCGCATCAAGAGCGGATGCTGTTGTT 640  
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288  
DB 641 AGCGTTGGCTAAATCGATCATTTTCAGATGTTAAATCAAGTAACCTCTGTTATTCCC 700  
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308  
DB 701 ACTGAAGCACTTAACAAAACCTTAGATGCTAACGGGTTTAAAGTTGGCGATGTTTA--- 757  
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328  
DB 758 -----ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGTAAAAGAGCAGCTATGCA 808  
QY 329 PheAspGlyArgTrrLeuAspArgSerValAspValIle-----LeuProAspAsn 345  
DB 809 AGTGTAGTGCCTAT---AAGCGCAACAGTTGAACCTATTGTCAATAGCTACCAAAATAT 865  
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
DB 866 CGCCTGAAATTTTAAATCAAAATCAATGAAGATGATAAAGCAAAATTTGGCATCATTAAC 925  
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
DB 926 TTC-----AAGGGAACGAA---TCTGTTAGTAGCAGCTACATTACAAGAA 967  
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402  
DB 968 CAAATGGAATTACAACCTGATTCTGTGGGAAATTTATGGGAAATAATTTGAAGTGCG 1027  
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
DB 1028 CAATTCGAGAAAGATTGCGAGTCAATTCGT-----GATTAT 1063  
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438  
DB 1064 TATTAAATATGGCTATGCCAAGCACAATTAATAACGGGATGTTTCAGCTAAATGAT 1123  
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456  
DB 1124 GAAAAACAAAAGTTAATGAACCATTTGATGTAATGAAGGTTTACAGTATGACCTTCGT 1183  
QY 457 -----GluSerThrLeuGluProValIleGlu 465  
DB 1184 AGTCGACCGCATTAAGTAATCTGGGAGGTATGCTGCCGAGCTTGAACCTTTACTTTCA 1243  
QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484  
DB 1244 GCATTACATTAAATGATACTTTCGCCCGCTAGTGATATTGCCAGATGAGAA----- 1294  
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504

DB 1295 ---AATGCAATTAAGCAAACTTGGAGAACCGGTTACGCTAGCGCAACCGTAAATTCA 1351  
QY 505 MetProasp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520  
DB 1352 GTACTGATTTTGTGATGCAAAATAAAACATTAGCGATAACCCCTGTTGTTGATCTCGGA 1411  
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533  
DB 1412 CGAGCTTTAACTGTTCCCAACTTCGGCTTTGAAGGAATAACCGTTTCTGCTGATAGCACT 1471  
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550  
DB 1472 TTACGTCAGGAAATCGCCCAACAAGAGGAACCTTGGTATTAATTCACAAATAGTTGAGTTA 1531  
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562  
DB 1532 GGAATAATTCGCTTAGATCGTACAGGTTTCTCGAAACAGTTCGAAACCGAATTCATCCT 1591  
QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574  
DB 1592 ATCAATGGTAGTATGATGAAGTGTGCTGATATATAAGTCAAA----- 1636  
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592  
DB 1637 -----GAACGTAAACACGGTAGTATCAACTTTTGTGTTTGTGTTACGGT 1678  
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608  
DB 1679 ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTAAACAGATAATTTCTTGGGAACAGG 1738  
QY 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
DB 1739 GCGCAGCTAAGTAGTAGCTGACGAAAAATGATTGTGTACGAGTGTCAATTTTCGGTTAT 1798  
QY 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638  
DB 1799 ACCGAGCCCTATTACTAAAGATGGTGAAGTCTTGGTGAATAATGTTTCTTGAANAAC 1858  
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658  
DB 1859 TACGATAACTCTAAAGTGATACATCTCTAACTATAAGCGTACGACTTACGGAAGTAAT 1918  
QY 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671  
DB 1919 GTTACTTTAGGTTTCCTGTAAATGAAATAACTCTATTATGTAGGATTAGGTCATACC 1978  
QY 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686  
DB 1979 TATAATAAATTAAGTAAGTTCCTGTCTAGAA-----TATAACCGTAATTTATATATAT 2029  
QY 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706  
DB 2030 CAATCAATGAAATTAAGGTAATGCGATTAAACAAAT----- 2068  
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnAlaLeuLeuAlaGlyVal 726  
DB 2069 -----GACTTTGATTTTCTTTTGTGTTGG----- 2092  
QY 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
DB 2093 -----AACTATACAGCCTTAATAGAGGCTATTTCCTCCAACT 2128  
QY 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762  
DB 2129 AAAGGGTTAAAGCAAGTCTTGGTGACAGCTTACTATTCCAGGTTCTCATACAATAATAC 2188  
QY 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
DB 2189 TCAAACTAAGTCAGATGTACAGGTTTCTACCCATTA----- 2227  
QY 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800

Db 2228 GACAGAGATCACCTCTGGTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGGTTTT 2287  
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArg 820  
Db 2288 GGAACAACAGCGTTTACCGTTCTATCAAACTTATACAGCGGTGGCATCGTTTCATTACGT 2347  
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 2348 GGTTCGTATTGTAGTAGTATGGAGCTAACGCAATTTATGCCGAATATGGTAATGGTAGT 2407  
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlnValLeu 842  
Db 2408 GGTACTGGTACTTTAAAGAGATAGTCTGTAT-----GTATGGTGGTAAATGCAATC 2461  
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856  
Db 2462 GCTACAGCTAGCGCAGAGTAAATTTGCGCAACTCCATTTGTGAGCGATAAGAGCAAAAT 2521  
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869  
Db 2522 ACGGTCCGAACCTCCTATTGTTGATCGGCAAGTGTTCGGAATACTAAATGGAATCA 2581  
Qy 870 -----AspLysGlyPheThrAsn 875  
Db 2582 GATAAAATGGATTAGAGCGCATGATTAAAGACATTCCTGTGATATGGCAATCAAGC 2641  
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrPalaSerProValGlyGlnValArg 895  
Db 2642 COTATTCGGCTCTACAGGTGCGATTCCCAATGCGCAATCTCTATTGGGCCATTGGTA 2701  
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913  
Db 2702 TTCCTTTATGCCAACCAATTAAATAATATGAAATATGATGTGCGAACAGTTCCAAATTT 2761  
Qy 914 PheIleGlyThrProPhe 919  
Db 2762 AGTATGGAGGTCTTTC 2779  
RESULT 19  
US-07-689-008-5  
Sequence 5, Application US/07689008  
Patent No. 5268274  
GENERAL INFORMATION:  
APPLICANT: Ben-Bassat, Arie  
APPLICANT: Calhoon, Roger D  
APPLICANT: Fear, Anna L  
APPLICANT: Gelfand, David H  
APPLICANT: Meade, James H  
APPLICANT: Tal, Ronny  
APPLICANT: Wong, Ring  
APPLICANT: Benzimid, Moshe  
TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE  
TITLE OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McCutchen, Doyle, Brown & Enersen  
STREET: Three Embarcadero Center  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/689,008  
FILING DATE: 19910422  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 337,194  
FILING DATE: 12-APR-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 496,236  
FILING DATE: 23-MAR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Lisabeth Feix  
REGISTRATION NUMBER: 31547  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 393-2000  
TELEFAX: (415) 393-2286  
TELEX: 340817 MACPAG SFO  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3957 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-689-008-5  
Alignment Scores:  
Pred. No.: 8.13e-07 Length: 3957  
Score: 159.50 Matches: 201  
Percent Similarity: 31.65% Conservative: 131  
Best Local Similarity: 19.16% Mismatches: 360  
Query Match: 3.37% Indels: 357  
DB: 1 Gaps: 51  
US-09-914-168-2 (1-919) x US-07-689-008-5 (1-3957)  
Qy 43 ProAlaHisAspThrAlaIle----- 49  
Db 490 CCGCGCATTCCTCGCGGTGGAATACTACCAGACCATGCGGGCGTGCAGCCAGTGG 549  
Qy 50 AsnGlnAlaLysAlaGlyAsnProValLeuLeuThr---ProGluGlnIleGlnAla 68  
Db 550 GACCAGGCCGCGCGGGCTTCGCGGTGCTGCTCAACCCGAGGATACCCGCC 609  
Qy 69 ArgLeu-----AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu--- 84  
Db 610 CAGCTCGCTTTCGCCAGGCCCTGACCTATAATACCTCAGCCGATGGAAGGCTGACC 669  
Qy 85 -----AspValValAsnPheAspGlnSerProIle---SerArgIleGlyGlu 100  
Db 670 CCGCTCAAGGATCTCCAGTCTTCGCGAGCCAGCGCGTCAATCTCGAGCCGCCCGCGCG 729  
Qy 101 GlnSerProLeuGlyLeuAsp---MetSerValIleGluGluThrThrProLeuSer 119  
Db 730 CAGTCTACCGCCAGACCCCTGAGCTGGCTGCGGTCAATCTCGAGACGAGCCCTCATG 789  
Qy 120 LeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIle 139  
Db 790 GAGCAGTGGCTTTCGCC-----CACCCCAATGAT----- 819  
Qy 140 ProGluTyrGlnGlyGluGlnProAsnSerGluValValValProProThrLeuGluPro 159  
Db 820 -----ACCGCGTGGCGAGCATATGCTCCACCCCGCGTGGTCCGCGG 864  
Qy 160 GluLysProGlyLeuIleLysArgLeuTyrAlaAtgLeuPheAsnAspGlyValAsnLys 179  
Db 865 GACAAGCGCGGTTCGCCGCGCAG----- 888  
Qy 180 ValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIle 199  
Db 889 -----GCAGTTTACCAGACGCTTACCGCGCGTTCCTGCGCGAGCC 930  
Qy 200 GlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIle 219  
Db 931 GAGCAGTCTTTCAG----- 945  
Qy 220 ThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu 239  
Db 946 -----TCGGCGTTGCAGATCAATTC----- 966

QY 240 ValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIle 259  
Db 967 -----CATGATGCTGAT-----TCGCTT 984  
QY 260 GlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAla 279  
Db 985 GGTGGCATGGGCTCGTAGCATGGCGAGGGCGATACCGCG-----GAGCGCGCGCGC 1038  
QY 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAlaGluVal 299  
Db 1039 TATTTTCAAGAGCGATGGCGCGCGACCCCAAG-----ACCGCGCATCGCTGG 1086  
QY 300 ProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIle 319  
Db 1087 CGCCCGCGCTTCGGGGCATGGCGCTACGCGCGAGTATGCTTCGGTTCCGTCAGTTGATT 1146  
QY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrp----- 333  
Db 1147 -----GCCGCCCATCAATATACCGAGGCGCAAGCAGCATGTCGCAGCTGGCC 1194  
QY 334 -----LeuAspArgSerVal 338  
Db 1195 CGCCAGCGCGCGCATATACTAGCGCGACCTCATGCTGGCGGACCTGCAGCGCTCGACC 1254  
QY 339 AspValIle----- 341  
Db 1255 GCCCAGATTGCCGCGCGCAGCAGGAATATCGTGCATCTCTGCGGTAGGCCCAATAAC 1314  
QY 342 -----LeuProAspAsnThrAlaAspVal 349  
Db 1315 CAGTTGGCCCTCATGGGCTGGCCGGGTAGACATGGCGCAGGGCAACACGGCGGAAGCA 1374  
QY 350 ---SerLeuIleTyrAspThrGlyThrGlnTyr-----ArgPheAspGluValValPhe 366  
Db 1375 CGCCAGCTCTGCTCGCTGTCGGCCCGCAATATGCAAGCCAGGTGGCGAGATGAGGTT 1434  
QY 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386  
Db 1435 TCGGCGCTGATGGCGGCTGCTCCAGACATCGGATTCAGCGCGCAGGTTTCCATCTG 1494  
QY 387 ArgGluLeuLeuGluGlnLeu-----LeuThrValAsnMetGlyGluAla 401  
Db 1495 CGCGAAGCGATGGCCAGGCGCCAGCGCTGACCCCTGGGTGGCGCATCAACCTTGCCAATGCG 1554  
QY 402 TyrAsnLeuGln-----AlaValArgAlaLeu----- 410  
Db 1555 CTGACGACGAGGCGAGGTGGCCGAGCGCGCGCTGATGACGCCATCTCTGGCGAAT 1614  
QY 411 -----SerAsnAsp 413  
Db 1615 CCCGTACCGCGCAGGACCGCGCGGTATCTTTTATACCTATGCTAGTGGCAATGAT 1674  
QY 414 LeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGlu 433  
Db 1675 CGCATGACCCCGCAGCTTCGGCTGCTGTCGCTCGGATTAATCTCTGCGCATCCGT 1734  
QY 434 GlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAla 453  
Db 1735 TCCATCGCC---GAGGAATGGAATCAAGCAGGATCTGGCCAGCGCGCTTCCATGTTG 1791  
QY 454 GlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIle 473  
Db 1792 TCCAAACCGGTGCGCTGATCCCGGAGGCGCTGACCCAGCGCTGATCCGAGCGCGCGC 1851  
QY 474 LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn 493  
Db 1852 GCGTGGCGGTGGCTGACCTGCTCCCGCAGCGTGGCGACATGGTGCATGCCCGCATGGCA 1911  
QY 494 LeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgVal----- 510  
Db 1912 CTGCGTATCCCGCTCGACGCGCGCATCATGATCTCTCGCGCCGACCGCGCTGCTCTATGCC 1971  
QY 511 -----LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIle 527

Db 1972 ACCGAATACATGAAGATCAGCAACCGGTCGGCTGGCGGG---CTGTCGCCCGCGTGG 2028  
QY 528 SerAspAlaVal---SerAlaValAlaArgAlaIleLeuProAspGluSerGlu----- 544  
Db 2029 GGGGATGGCACGGCTCGGCTACAGGAAGCGGTTGCTGCCGAGCAGCTGCACAGCGCTC 2088  
QY 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561  
Db 2089 CAGCAACTGGCGCATGGGATCTCGCTGGCGAGTCGCTCAACACCGCTGGCGAC 2148  
QY 562 ProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLys 581  
Db 2149 CAGGCGCAGCGCTATGATCATCTGCCCGCGCTGCAGCGCGACCCCGGAGCGGCATCG 2208  
QY 582 ProArg-----AspGlyGlnIleGlyLeuGlyTyrGly 592  
Db 2209 CCCAAGCTGGGCTGGCGCGCTGTATATATGGCACGCAAGCGCGCAAGCGCTCGAG 2268  
QY 593 SerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGly 612  
Db 2269 ATCGAC-----CTTGGCGTGTGGCGCGCAACACCGCAGGACCTTTCATGCG 2313  
QY 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyVal-----LysLeu 630  
Db 2314 CGACAGGCTGGCTGCGCGCGCTCAACAGCAGCACAACACAGCTTGGCCACCGCTT 2373  
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650  
Db 2374 GCCATGATGGCTGCAGGAAGCCGATGGATGCCCTGGCTGGCTGGCGCATGGCGGTG 2433  
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670  
Db 2434 GCTGACCGCGCGATGGCCAC-----GGCAGCGCAGCATCGAG 2472  
QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyTyrTrpAsnArgThrTyrSerLeuArg 690  
Db 2473 -----GATCTGGCGCGCGCTATGACCTGCGC 2499  
QY 691 TyrArgLeu-----AspLys- 695  
Db 2500 CTGACGAGCTCGAGGCGCGCGCGCTGTCGGCGCGGTGCTGGCGAGGAGATGCG 2559  
QY 696 LeuLysThrGlnAlaProProGluThr-----Trp 705  
Db 2560 CTT-----GCTCGCGCTCGACCAACCGCTTCGCGCGCGTGGCTACGCCACCAG 2610  
QY 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAla----- 721  
Db 2611 ACGGAACCTGGCGCGCTGTGACCGGTGCTTACAGCGCGCGGCGCATCGCCCGAT 2670  
QY 722 -----LeuLeuAlaGlyValAlaValHis---LysThrValAlaAspAsnLeu 736  
Db 2671 ACCTCGGACCATGCTCTCTCTCCATCGCAGCGCAGATCCGACGCTGCTGAGAACCTT 2730  
QY 737 -----ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753  
Db 2731 GCCCTTCCATCATGCTGGCTCGCGGTTCCGCTCGCTG---TCGGGTGAGCATGGC--- 2784  
QY 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAla-----ArgAlaGly 768  
Db 2785 ATGGCGCGCTGACGGAAGCAACATTCCTCATGTCGGCGCGCTGCGCTGACAGCGCGT 2844  
QY 769 IleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMet 788  
Db 2845 GCTTCGCC--- 2856  
QY 789 ThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsn----- 803  
Db 2857 ACCTTCTCATCAGCCCAACCATGATCTCGTGGCAACCTCAACACCGGGTTCCTGCTAT 2916  
QY 804 HisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyr--- 822

Db	2917	GATGTCGC	-----CGTTAT-----GGCAGATGATGGCGTGCAGGCATATAAC	2961
Qy	823	AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGly	-----839	
Db	2962	CAGTACGATAGC	-----TATACCAACGCGGCGAGGACCAG	2997
Qy	840	GlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeu	859	
Db	2998	CAGCGCATCGCGCTGCACGCGCCAG	-----3024	
Qy	860	AlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIle	879	
Db	3025	-----	-----GCCGGGTTGGCCCGGATGTCAGTTT	3051
Qy	880	Gly	-----AlaGlyValGlyValArgTTPAlaSerProValGlyGlnVal	894
Db	3052	GGCAATAGCTGGTGGCGCCGATGTGGT	-----GCGTCGCCCATCGGCTTCCCC	3102
Qy	895	ArgValAspValAlaThrGlyValLys	903	
Db	3103	ATCACCAAGCTGCTGGCGGCTGTCGAG	3129	
RESULT 20				
US-07-689-008-1				
: Sequence 1, Application US/07689008				
: Patent No. 5268274				
: GENERAL INFORMATION:				
:	:	APPLICANT: Ben-Bassat, Arie	:	:
:	:	APPLICANT: Calhoon, Roger D	:	:
:	:	APPLICANT: Fear, Anna L	:	:
:	:	APPLICANT: Gelfand, David H	:	:
:	:	APPLICANT: Meade, James H	:	:
:	:	APPLICANT: Tal, Roni	:	:
:	:	APPLICANT: Wong, Hing	:	:
:	:	APPLICANT: Benziman, Moshe	:	:
:	:	TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE	:	:
:	:	TITLE OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON	:	:
:	:	NUMBER OF SEQUENCES: 15	:	:
:	:	CORRESPONDENCE ADDRESS:	:	:
:	:	ADDRESSEE: McCutchen, Doyle, Brown & Enersen	:	:
:	:	STREET: Three Embarcadero Center	:	:
:	:	CITY: San Francisco	:	:
:	:	STATE: California	:	:
:	:	COUNTRY: USA	:	:
:	:	ZIP: 94111	:	:
:	:	COMPUTER READABLE FORM:	:	:
:	:	MEDIUM TYPE: Floppy disk	:	:
:	:	COMPUTER: IBM PC compatible	:	:
:	:	OPERATING SYSTEM: PC-DOS/MS-DOS	:	:
:	:	SOFTWARE: PatentIn Release #1.24	:	:
:	:	CURRENT APPLICATION DATA:	:	:
:	:	APPLICATION NUMBER: US/07/689,008	:	:
:	:	FILING DATE: 19910422	:	:
:	:	CLASSIFICATION: 435	:	:
:	:	PRIOR APPLICATION DATA:	:	:
:	:	APPLICATION NUMBER: US 337,194	:	:
:	:	FILING DATE: 12-APR-1989	:	:
:	:	PRIOR APPLICATION DATA:	:	:
:	:	APPLICATION NUMBER: US 496,236	:	:
:	:	FILING DATE: 23-MAR-1990	:	:
:	:	ATTORNEY/AGENT INFORMATION:	:	:
:	:	NAME: Murphy, Lisabeth Feix	:	:
:	:	REGISTRATION NUMBER: 31547	:	:
:	:	TELECOMMUNICATION INFORMATION:	:	:
:	:	TELEPHONE: (415) 393-2000	:	:
:	:	TELEFAX: (415) 393-2286	:	:
:	:	TELEX: 340817 MACPAG SFO	:	:
:	:	INFORMATION FOR SEQ ID NO: 1:	:	:
:	:	SEQUENCE CHARACTERISTICS:	:	:
:	:	LENGTH: 9540 base pairs	:	:
:	:	TYPE: NUCLEIC ACID	:	:
:	:	STRANDEDNESS: single	:	:
:	:	TOPOLOGY: linear	:	:

QY 300 ProLeuLeuLeuGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuile 319  
Db 6091 CGCCGCGCGTGGCGCATGCGCGTGCAGCGCGAGTATGCTCCGTTCCGCCATGATT 6150  
QY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrp 333  
Db 6151 -----GCCGCCCATCAATATACGAGGCCAAGCAGCAGTTCGACGCTGGCC 6198  
QY 334 -----LeuAspArgSerVal 338  
Db 6199 CGCAGCGCGCGCAGTATCTGGCGCGACCCCTCATGCTGGCCGACCTGCAGCGCTGACCC 6258  
QY 339 AspValile 341  
Db 6259 GGCAGATTCCGCGCGCAGCAGGAATATCGTGGCATCCTGTCGCGTGAGCCCAATAAC 6318  
QY 342 -----LeuProAspAsnThrAlaAspVal 349  
Db 6319 CAGTTGGCCCTCATGGGCGTGGCGCGGTAGACATGCGCGAGGCCGACACGCGGGAAGCA 6378  
QY 350 ---SerLeuileTyrAspThrGlyThrGlnTyr-----ArgPheAspGluValValPhe 366  
Db 6379 CGCCAGCTCCTGTCGCGTGTGCGCGCCCATATCCAAAGCCAGGTGGCGGAGATCGAGTT 6438  
QY 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386  
Db 6439 TCGGCGCTGATGGCGGTGCGTCCAGACATCGGATTTCAGCGCGCAAGGTTTCATCCTG 6498  
QY 387 ArgGluLeuLeuGluGlnLeu-----AlaValArgAlaLeu 401  
Db 6499 CGGAAAGCATGGCCAGGCCCCACGTCGACCCCTGGGTGGCGCATCAACCTTGCCAATGCG 6558  
QY 402 TyrAsnLeuGln-----AlaValArgAlaLeu 410  
Db 6559 CTGCAGCAGCAGCGGACGTGGCGGAGCGCGCGGTGATGCGGCCATCCTGGCCAAAT 6618  
QY 411 -----SerAsnAsp 413  
Db 6619 CCGGTACCCGCGCAGGACCGCCAGCGGTATCCTTTATACCTATGTTAGTGGCAATGAT 6678  
QY 414 LeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluAtrgGlu 433  
Db 6679 CGCATGACCCGCGCAGCTTCTGGTGTGTCGCGCGGATATTCTCTCGCATCCGT 6738  
QY 434 GlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAla 453  
Db 6739 TCCATCGCC---GAGGAATGGAAATCAAGCAGGATCTGGCCAGCGCCTGTCCATGGTG 6795  
QY 454 GlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIle 473  
Db 6796 TCCAACCCGTCGCTGTATCCGCGAGGCGCTGACCCAGCCTGATCCGACCGCGCGCGC 6855  
QY 474 LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn 493  
Db 6856 GCGTGGCGTGGCTGACCTGTCGCCCGCAGCGTGGGACATGGTGATGCCCGCCATGGCA 6915  
QY 494 LeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgVal-----510  
Db 6916 CTGCTATCGCTCGAGCGCCACCATCGATCTCTCGCCCGCAGCAGCGCTGTCTATGCC 6975  
QY 511 -----LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIle 527  
Db 6976 ACCGAATACATGAATACAGCAACCGCGTGGCGCGTGGCGG---CTGCTGGCGCGCGTG 7032  
QY 528 SerAspAlaVal---SerAlaValAlaArgAlaIleLeuProAspGluSerGlu-----544  
Db 7033 GGGATGGCAGCGGCTGCGGTACAGGAAGCGCGTGTGCGCCGAGCAGGTGCGAGCGCTC 7092  
QY 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561  
Db 7093 CAGCAACTGGCGCATCTCGGTGGCGCATCTCGGTGGCGCATCTGCTCAACCAAGCGTGGCGAC 7152

QY 562 ProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaLaserAspLys 581  
Db 7153 CAGCGCAGCGCCTATGATCATCTGCGCCCGCTCAGCGCCACCGGAGCGGCACATCG 7212  
QY 582 ProArg-----AspGlyGlnIleGlyLeuGlyTyrGly 592  
Db 7213 CCCAAGCTGGCGCTCGCGCGCTGTATATGTCACGGAAGCGGCAAGCGCTCGAG 7272  
QY 593 SerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGly 612  
Db 7273 ATCGAC-----CTTGGCGTCTCGCCACAAACCCGAGGACCTTGATGCG 7317  
QY 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyVal-----LysLeu 630  
Db 7318 CGCAGGCTCGCGTGCAGCGCGGTCAACAGCGACCAACAGCCTTGCACCCGCTT 7377  
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650  
Db 7378 GCCATGGATGGCTGCAGGAAGCCCGATGGATGCCCTGCTGGTGGCATGGCGGTG 7437  
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670  
Db 7438 GCTGACCGCGCCATGCCAC-----GGCGAGCGCCACCATCGAG 7476  
QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArg 690  
Db 7477 -----GATCTGCGCGCGCTATGATGACCTGCGC 7503  
QY 691 TyrArgLeu-----AspLys 695  
Db 7504 CTCACAGGTTCGAGGCGCAGCGCGCGCTCTGCGCGCGGTCTGCGCAGGAAGATGCG 7563  
QY 696 LeuLysThrGlnAlaProGluThr-----Trp 705  
Db 7564 CTT-----GCTCGCGCTCGACCAACCCGCTTCCGCGCGCTGCTACGCGCCACAG 7614  
QY 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAla-----721  
Db 7615 ACGGAACCTTGGCGCGCTGTGACCGGTGGCTCTACAGCGCGCGAGCGCATCGCCGAT 7674  
QY 722 -----LeuLeuAlaGlyValAlaValHis---LysThrValAlaAspAsnLeu 736  
Db 7675 ACCTCGGACCATGCTCTCTCCATCGCAGCGCCACATCGCACGCTGCTGAGAACCTT 7734  
QY 737 -----ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753  
Db 7735 GCCCTTCCATCGATGGTGGCGCTCGCGTCCGCTCCGT---TCGGGTGAGCATGCG---7788  
QY 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAla-----ArgAlaGly 768  
Db 7789 ATGGGCGCGCTGACGGAAGCAACATTCCTCCATCGTGGCGCGCTGCGCGTGCAGCGCGT 7848  
QY 769 IleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMet 788  
Db 7849 GCTTCGCC-----CTG 7860  
QY 789 ThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsn-----803  
Db 7861 ACCTTCTCGATCAGCCCAACCATGATCTGTGCGGCAACCTCAACACGGGTCCGCTCAT 7920  
QY 804 HisValProTyrArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyr---822  
Db 7921 GATGTGCG---GCTTAT-----GGCAGCATGATGGCGGTGAGGATATAAC 7965  
QY 823 AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGly-----839  
Db 7966 CAGTACGATAGC-----TATACCAACCGCGCGAGGACCCAG 8001  
QY 840 GlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeu 859  
Db 8002 CAGCGCATCGCGCTGCGCAGCGCGAG-----8028  
QY 860 AlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrIle 879







Qy 272 oValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys----- 290  
Db 23402 TCAGCTTGAGCAGATACGCGCTTGGGAGCAGGGCGGTCCAGTGATGTCGACGGAT 23461  
Qy 291 -AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHisG1 310  
Db 23462 AGCCATCTCTGCATGCGCAGGAGAAACCGGTGCGCGTACTGCGGATATT----- 23510  
Qy 310 yLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAs 330  
Db 23511 -----GATAATAACCTCAATGCCCTTGGCAGTACGCTGAAGTATCTGTCT 23554  
Qy 330 pGlyArgTyr-----LeuAs 335  
Db 23555 TGATTTCTGGAGTCGTTCTGGGATCGGCCATGAATATTGGTGTGAAGACTCGCTGGA 23614  
Qy 335 pArgSerValAspValIle-----LeuProAspAsnTh 346  
Db 23615 TGAACAGATTGCCGCTTTACAGAGAGAAGTGTGCGGGCGGAAAAAGACTCCCTGGACGGC 23674  
Qy 346 rAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPh 366  
Db 23675 ATCATCTCTCAGGTTGAATACGAT-----CAGCAGGCTCTTACGAT----- 23717  
Qy 366 ePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLy 386  
Db 23718 -----CTTCAGGAGAAAAAACCCAGAGGATTTGCGAGTGCATAAAGAGCAGGCAGA 23770  
Qy 386 sArgGluLeuGluGln-----LeuLeuThrVa 396  
Db 23771 CGGGAATATACAGGAGCAACAGAACCGCGTAATGCTGAAATGCTGCATGAAACCGGAT 23830  
Qy 396 lAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAl 416  
Db 23831 GAATGAAACGGAACAGCAGACATCAGCGTGAATACGCGTATTAT-----GC 23881  
Qy 416 aThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleG1 436  
Db 23882 CATGCAGTACGCGCATCAGGCTGTTCAGGATGCGCG-----ATACA 23923  
Qy 436 nAsnAspGlnValSerPheGluGln-----SerSerSerSerArgThrGluProAl 453  
Db 23924 ACGTGAATATGAACGTTACGAGAACCCCTGGCATCCCGTAAGAAAAAACACCGCGAAAC 23983  
Qy 453 aGlnValAspGluSerThr-----LeuG1 461  
Db 23984 CCGTAATGATGAGGCCACCCCGTTATTGCTCAGTACAGTCAGCAACAGGCACAGGTGGA 24043  
Qy 461 uProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleG1 481  
Db 24044 AGGACAGATTGCTGCTGCCACAGACAGTCAGCGCATT----- 24080  
Qy 481 uPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgH1 501  
Db 24081 -----GCCACGGAAGGATGACAGAGCGCGTAAACAGCTTCGCTGCGACGACGCG-- 24134  
Qy 501 sLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnAr 521  
Db 24135 -ATCAGCGACTGGACGGGAAAAAATCAGCGCA-----CATGAA-----AA 24175  
Qy 521 gSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAs 541  
Db 24176 GAGTGTCTGGCCGCTAAAGATGAATTCAGTTCAGGCACTGACG-----CTGCTGGA 24226  
Qy 541 pGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysTh 561  
Db 24227 TGTAACACAGCAGGAG-----CTTCAGAAACACAGCGCACTCAACGAGCTGAAG-- 24275  
Qy 561 rProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLy 581  
Db 24276 -----AAAAAACAAATTCAGCTGACCCAGTCACTGCGTGAAGAGA 24316

Qy 581 sProArgAspGlyGln-----IleGlyLeuGlyTyrGlyLe 593  
Db 24317 CGCGCTCAGCGTCAGCAACATACCTGGATATGCCACGCGTGGGTATG----- 24365  
Qy 593 rAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTy 613  
Db 24366 -----GGTGATCAGCAGCAGCAGCGATATCAGGTACAACCTGAGTCTTCGCGAATA 24418  
Qy 613 rGlnAlaGlyAlaGlu-----LeuArgLeuSerGluAspLysLysGlyValLysLeuTy-- 631  
Db 24419 CCACACACGCTGAGCAGCTGAGCGGGATAGTGAGCAGAAAGGACA-----TATAA 24472  
Qy 632 -----AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuAr 645  
Db 24473 CACGGATGACTACAGAAAGCGCAGCAGCGCTGACGAGAGCTGAAACCGACAACATAA 24532  
Qy 645 gAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSe 665  
Db 24533 TGAGAAATCGCGTTACTGTGGCAACAGCTT----- 24560  
Qy 665 rThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnAr 685  
Db 24561 -----GAAGTTGTACAG-----GGTAACCTGG--AA 24583  
Qy 685 gThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTr 705  
Db 24584 AAACGGAGTCTCGCT-----GCATT 24604  
Qy 705 pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaG1 725  
Db 24605 TCAGGATTTTACCCTGGATGAGATATACGGCAGAAACAGACAGAA----- 24650  
Qy 725 yValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgG1 745  
Db 24651 ---CAGGTGTCTCGTCAGCGCTTCAGCAACATGGGAATGCCCTGGCAACTTTTGTCTAC 24706  
Qy 745 nArgTyrSerLeuGluValGlySer-----SerGlyLeuValSerAspAlaAsnMetAl 763  
Db 24707 TACCGGCAAACTCAATTTCAAATCCTTCACCTCTCTGTGTGTCAGATATATGCGGAAA 24766  
Qy 763 aIleAlaArgAla-----GlyIleSerGlyValTyrSerPheG1 776  
Db 24767 CCTGGCGCAGGACCACTGATGAAATCGAATAAAGGGATTCGAGTGTACTGGGATTGGA 24826  
Qy 776 yAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTy 796  
Db 24827 TCTCAGCAGCGCTTCCCTGTAAT-----GCCAATGGGGGATTTATCAGTCTGC 24874  
Qy 796 rIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla----- 813  
Db 24875 TGATTTGAGTCGTTACAGTGGCAGCGGTGTTAACCGTCCGACGTTTTTTGCTTTTGCAAA 24934  
Qy 814 -----GlyGlyAspGlnSerIle-----ArgGlyTyAl 823  
Db 24935 AGGCGCGGTGTGATGGGGGAGCGGACCTGAAGCCATTCTGCCATTGCGTGGTGGC 24994  
Qy 823 aHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAl 843  
Db 24995 TGACGGTAAGCTGGGGGTGTGGCGGAT-----ATTGGGGGTTCAGGTATGCG 25042  
Qy 843 aValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheG1 863  
Db 25043 GATGTTTTTCCCGCGGATACAACATCGAGATCAATAACAT-----GGCACGAACGG 25093  
Qy 863 yAspIleGlyAsnAla-----TyrAspLysGly 872  
Db 25094 GCAGATAGTCCGCTGCCCTGAAGCGGTTTATGACCTCGGG 25136

## RESULT 23

US-08-621-944A-2  
; Sequence 2, Application US/08621944A  
; Patent No. 6440425

; GENERAL INFORMATION:

APPLICANT: SASAKI, Ken  
 APPLICANT: HARKNESS, Robin E.  
 APPLICANT: LOOSMORE, Sheena M.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER  
 MEMBRANE PROTEIN OF MORAXELLA  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/621,944A  
 FILING DATE: 26-MAR-1996  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/478,370  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-587  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5976 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-621-944A-2

Alignment Scores:  
 Pred. No.: 0.000655 Length: 5976  
 Score: 136.00 Matches: 191  
 Percent Similarity: 32.95% Conservative: 128  
 Best Local Similarity: 19.73% Mismatches: 351  
 Query Match: 2.88% Indels: 298  
 DB: 4 Gaps: 49

US-09-914-168-2 (1-919) x US-08-621-944A-2 (1-5976)  
 Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHis 41  
 Db 1357 CCATATTGGATAAAACAACTTAAGTGGGTAGTGTGCAATTACCATAGACAATGGC 1416  
 Qy 42 ValProAlaHisaspThrAlaIle---AsnGlnAlaLysAlaGlyAsnProValLeu 60  
 Db 1417 ATTGATCGAGTATAAAAGATCAGTAATCTTGCCAAAGGTAGCAGTCTACGATGCG 1476  
 Qy 61 LeuThrProGluGlnIleGlnAla-----ArgLeuAsnAla---AlaGlyLeuAsn 76  
 Db 1477 GTTACCATCCAGCAGCTCAAGCGCCAGCCAGCTTCTTAACGCGAGCGCTGGCATCAGT 1536  
 Qy 77 AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAsnAspGlnSerProIleSer 96  
 Db 1537 GTCACACCTACTGAAATATACAGTTGATGCTAAGAGTGGCAATGTTACCGCCGCCAACTTAC 1596  
 Qy 97 ArgIleGlyGlnSerProProLeuGlyLeuAsp-----MetSerVal 111  
 Db 1597 AACATTGGCGTGAACACCCAGCTTAACAGTGGCACTAGTGATGATAAAATTAGTGT 1656  
 Qy 112 IleGluGluThrThrProLeuSerLeu-----GluGluLeuPheAlaGlnGluSerThr 129

Db 1657 AAGGGTAGTGGTACGAACAATAGCTTAGTTACCGCGCAACATTTTGGCAGCTATCTAAT 1716  
 Qy 130 GluMetGlyIleAsnProAsnAsp-----TyrIleProGluTyrGlnGly 144  
 Db 1717 GAAGTCAATCGAACGGCTGACAGAGTCTCTACAAAGCTTTTACCGGTTAAAGAAGACAGCAT 1776  
 Qy 145 GluGlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeu 164  
 Db 1777 GATGACGCCAACGCTATCACCGTGGCTAAAGATACGACAAA----- 1818  
 Qy 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys 184  
 Db 1819 -----AATCGCGCGCAGTCAGCATCTTAAACTCAAA 1851  
 Qy 185 AlaLys-----PheTyrGlnSer 190  
 Db 1852 GGTAACAAACGGCTTAACGGTGTCTACCAAAAAGATGGTACGGTTACCTTTGGGCTTAGC 1911  
 Qy 191 SerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla 210  
 Db 1912 CAAGATAGCGGTCTGACC-----ATTGGCAAAAGACCCCTAAACACAGTGGC---TTG 1962  
 Qy 211 AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla-----Met 225  
 Db 1963 ACTGTTAAAGATACCAACGACAAATC---CAAGTCGGTCTAATGGCATTAATTTACT 2019  
 Qy 226 AspLeuAsnGlySerIlePro-----ArgLeuArgGlnThrAlaLeuValAlaAlaArg 243  
 Db 2020 AATGTGAATGGTAGTAATCAGGTACTGGCATTTGCAAAATACCGCTGCATTACACAGAT 2079  
 Qy 244 AlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAsp 263  
 Db 2080 AAAATTGGCTTTGCTGTTCTGATGGTGGCAGTT----- 2112  
 Qy 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal--- 282  
 Db 2113 -----CATACAAACAAACCT---TATCTTGATCAAGACAAAGTACAGTTGGC 2157  
 Qy 283 -----ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296  
 Db 2158 AATGTTAAGATTACCAACACATGGCATTAACGCGGTGGTAAAGCCATCACAGGCTGTCC 2217  
 Qy 297 AspGluValProLeuLeu----- 302  
 Db 2218 CCAACACATGCTCTAGCATTTGCCGATCAAAGTACCCGCAACATAGAACTGGCCAAATCAATC 2277  
 Qy 303 -----IleGlyAspValPheHisGly---Lys 311  
 Db 2278 CAAGACAAGACAATAATCCAAACGCTGCCAGCATTAATGATATATTAATACAGCTTTAAC 2337  
 Qy 312 TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331  
 Db 2338 CTAAAAATAATAACACCCCATTCGCTTCTCC----- 2373  
 Qy 332 ArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsnThrAlaAspVal 349  
 Db 2374 -----ACTATGACATTTGTTGCTTGTGCAATGGCAATGCCACCCACCC 2418  
 Qy 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe---PheThr 368  
 Db 2419 ACAGTAACCCATGATACCGCTAAC-----AAAACCAAGTAAAGTGTATATGATGTGAAT 2472  
 Qy 369 IleAspProLysThrAsnGlnLeu-----PheThrAspProAspLysLeuProValLys 386  
 Db 2473 GTGGATGATACACCATTCATCTAACAGGCACTGATGACATAAAAACTTGGCGCTCAAA 2532  
 Qy 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuAla 406  
 Db 2533 ACCACCAACTGAAC----- 2547  
 Qy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr 425

Db 2548 -----AAAACAAGTCCTAATGTTAATACAGCACTAACTTTAATGTTAACTCTAGTGATGAA 2604  
Qy 426 GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln 444  
Db 2605 GATGCCCTTGTAAACCCCAAGACATCGCCGAAATCTAAACACCTAGCCCAAGAAATT 2664  
Qy 445 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464  
Db 2665 CACACCACCAAGGACACACACACCGCCCTACAAACCTTTACCGTTAAAAAGGTAGAT 2724  
Qy 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484  
Db 2725 GAAATAATAATGCTGATGACGCCCAACGCCATCAACCGTGGGT-----CANAAGAACCCA 2778  
Qy 485 SerAsnLeuIleGlnAspLysLeuValAlaAlaLysAlaArgHisLeuTyArg 504  
Db 2779 AATAAT-----CAAGTCAACACCTTAACACCTCAACAGGTCAAAACCGTCTTAAT 2826  
Qy 505 MetProAspArg-----ValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 2827 ATTAACCGCACAATAATGTTAGCTTACCTTTGTCATTAACACCAACAGCGGTCTTAA 2886  
Qy 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540  
Db 2887 -----GCGGCAAAAGCACCCCTAAAC 2907  
Qy 541 AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560  
Db 2908 GACGGTGGC-----TTGCTATTAAAAACCCCACTGGTAGCGAACAATCCAA 2955  
Qy 561 ThrProAlaAspValTyGlnSerLysValProLeuTyArgValPheValAlaSerAsp 580  
Db 2956 GTCGGTGTGATGGGTGAAGTTTGCCTCAAGTT----- 2988  
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600  
Db 2989 ---ATAATAATGGTGTGTAGTGCTGGC-----ATTGATGGCACAACCTGCATT 3036  
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyArgGlnAlaGluLeuArg 620  
Db 3037 ACCAGATGAATGGCTTTACTGGGACTAATGGC----- 3072  
Qy 621 LeuSerGluAspLysGlyValLysLeuTyArgAlaThrLysProLeuSerHisProLeu 640  
Db 3073 ---TCACCTGTATAAA-----AGCAAAACC----- 3093  
Qy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyArgGlnGluValPheGlyHis---SerThr 659  
Db 3094 -----CACCTAAGCAAA 3105  
Qy 660 AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGln 679  
Db 3106 GACGGCATTAAACCGCAGGTGTAAAGATTAACCAACATTCAATCAGGTGAGATTGCCCAA 3165  
Qy 680 Asn-----GlyGlyTrpAsnArgThrTyArgSerLeuArgTyArgLeu 693  
Db 3166 AACAGCCATGATGCTGTGACAGCGGC-----AAGATTATGATTAAAAACCCGAACCT 3219  
Qy 694 AspLysLeuLysThrGlnAlaProGluThrTrpGlnAspLeuProValAspPheVal 713  
Db 3220 GAA----- 3222  
Qy 714 AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal 732  
Db 3223 ---AACAAATCAGCAGTACTGCCAAAACAGCACAAACCTCATACAGAAATCTCAGTA 3279  
Qy 733 AlaAsp-----AsnLeuValAsnProMetArgGlyTyArgGlnArg 746  
Db 3280 GCAGATGAACAGGTAATAAATACTTACGGTTAGTAACCCCTTACTCCAGTTATGACACCTCA 3339  
Qy 747 TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet 762  
Db 3340 AAGACCTCTGTCTATCATCACCCTTTCAGGTGAAACGGGATTACCACCAAGGTAATAAA 3399

Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTySerPheGlyAspAsnAlaTyArgGlySer 782  
Db 3400 GGTGTGGTGGTGGGCATT-----GACCAAAACCAAGGCCTTA 3438  
Qy 783 AsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyArgIleTrpSerAspAsn 802  
Db 3439 ACCAGCCCTAAGCTGACCGTGGTAAATAATAATGCAAGGCAATTCATTGACAGCCAA 3498  
Qy 803 AsnHisValProTyArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyArg 822  
Db 3499 AAT-----GGTCAAAATACCATCACAGCACTA 3525  
Qy 823 AlaHisAspSerLeuSerProIleSerAspLysGlyTyArgLeuThrGlyGlnValLeu 842  
Db 3526 ACACACACTCTAGCTAATGTTACCAATGATAAGGT----- 3561  
Qy 843 AlaValGlyThrAlaGluTyArgAsnTyArgGluPheMetLysAsp-----LeuArgLeu 859  
Db 3562 AGCGTACGCACACACAGACAGGCGC---AATATAATCAAAAGACGACGACAAACCCGTGCC 3618  
Qy 860 AlaValPheGlyAspIle-----GlyAsnAlaTyArg 869  
Db 3619 GCACGATTGTTGATGTGCTAAGCCAGCGCTTTAACTTCAAGGCAATGCTGAAGCGTT 3678  
Qy 870 AspLysGlyPheThrAsnAspThr 877  
Db 3679 GACTTGTCTCCACTTATGACACC 3702

## RESULT 24

US-08-945-567D-2.  
; Sequence 2, Application US/08945567D  
; Patent No. 648386  
; GENERAL INFORMATION:  
; APPLICANT: SASAKI, Ken  
; APPLICANT: HARKNESS, Robin E.  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: CHONG, Pele  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF  
; TITLE OF INVENTION: MORAXELLA  
; FILE REFERENCE: 1038-745 MIS  
; CURRENT APPLICATION NUMBER: US/08/945,567D  
; CURRENT FILING DATE: 1996-04-29  
; PRIOR APPLICATION NUMBER: 08/431,718  
; PRIOR FILING DATE: 1995-05-01  
; PRIOR APPLICATION NUMBER: 08/478,370  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: 08/621,944  
; PRIOR FILING DATE: 1996-03-26  
; PRIOR APPLICATION NUMBER: PCT/CA96/00264  
; PRIOR FILING DATE: 1996-04-29  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 5976  
; TYPE: DNA  
; ORGANISM: Moraxella catarrhalis  
US-08-945-567D-2

Alignment Scores:  
Pred. No.: 0.000655 Length: 5976  
Score: 136.00 Matches: 191  
Percent Similarity: 32.95% Conservative: 128  
Best Local Similarity: 19.73% Mismatches: 351  
Query Match: 2.88% Indels: 298  
DB: 4 Gaps: 49

US-09-914-168-2 (1-919) x US-08-945-567D-2 (1-5976)

Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIleIleAsnHis 41  
Db 1357 CCATATTGGATAAAAAACAACCTTAAAGTGGGTAGTGTTCGAATTACCATAGACAATGCG 1416

Qy	42	ValProAlaHisAspThrAlaIle---	AsnGlnAlaLysAlaGlyAsnProValLeu	60
Db	1417	ATTGATCGAGGTAAATAAAGATCAGTATCTTGC	AAAGAGGTACGAGTGCTTAACGATGCG	1476
Qy	61	LeuThrProGluGlnIleGlnAla-----	ArgLeuAsnAla---AlaGlyLeuAsn	76
Db	1477	GTTACCATCAAGACAGCTCAAGCCGCAAGCC	TACTTTAAACGCGAGCGCTGCATCAGT	1536
Qy	77	AlaLysProGlnSerGlnAlaLysValValAsn	PhenPheAspGlnSerProIleSer	96
Db	1537	GTCAACCTACTCAAAATATACAGTTGATGCT	TAGAGATGGCAATGTTACGCCGCCAACCTTAC	1596
Qy	97	ArgIleGlyGluGlnSerProLeuGlyLeuAsp-	-----MetSerVal	111
Db	1597	AACATTTGGCGTGAACACCCAGCAGCTTAAC	AGTGTGACAGTGTGCACGTAGTATAAATTTAGT	1656
Qy	112	IleGluThrThrProLeuSerLeu-----	GluGluLeuPheAlaGlnGluSerThr	129
Db	1657	AAGGTTAGGTACGAACAATAGCTTAGTTAC	CGCGCAACATTTGGCGAAGCTATCTAAAT	1716
Qy	130	GluMetGlyIleAsnProAsnAsp-----	TyrIleProGluTyrGlnGly	144
Db	1717	GAACTCAATCGAACCGCTCACAGTCTCTACA	AAAGCTTTACCGTTAAAGAAAGACGAGT	1776
Qy	145	GluGlnProAsnSerCluValValValProPro	ThrLeuGluProGluLysProGlyLeu	164
Db	1777	GATGACGCCACGCTATCATCCGTGGCTAAG	ATACGACAA-----	1818
Qy	165	IleLysArgLeuTyrAlaArgLeuPheAsnAsp	GlyValAsnLysValProArgLeuLys	184
Db	1819	-----AATCCGCGCAGTCACCATCTCTTAA	AAACTCAAA	1851
Qy	185	AlaLys-----	-----PheTyrGlnSer	190
Db	1852	GGTAAAAACGGTCTAACGGTTGCTACCAAAA	AGATGTCAGGTTACTTTGGCGTTAGC	1911
Qy	191	SerGlnSerSerGlyThrSerAlaIleGlySer	HisGlnLysThrGluProTyrAla	210
Db	1912	CAAGATAGCGGCTGACC-----ATTGGCA	AAAGCACCCCTAAACACGATGCG---TTG	1962
Qy	211	AsnIleLysAlaAlaLeuGluAspIleThrGln	SerAla-----Met	225
Db	1963	ACTGTTTAAAGATACCAACGACGAACAATC	---CAAGTCGGTCTAATGGCATTAATTTACT	2019
Qy	226	AspLeuAsnGlySerIlePro-----	ArgLeuArgGlnThrAlaLeuValAlaAlaArg	243
Db	2020	AAATGTGAATGGTAGTAATCCAGGTACTG	CGCATTCGCAATACCCCTCGCATTACCGAGAT	2079
Qy	244	AlaValGlyTyrTyrAspIleAspLeuSerIle	IleArgAsnSerIleGlyGluValAsp	263
Db	2080	AAAAATGGCTTGTGGTTCCTGATGGTGCAGT	-----	2112
Qy	264	ValIleIleHisAspLeuGlyGluProValTyr	IleAspTyrArgAlaValGluVal---	282
Db	2113	-----GATCAACAACAACCT---TATCT	TGTATCAAGACAAGCTACAAGTTGCG	2157
Qy	283	-----ArgGlyGluGlyAlaAspAspLysAla	PhenThrThrValAla	296
Db	2158	AATGTTAAGATTACCAACACTGGCATTTAAC	GCAGGTGGTAAAGCCATCATCAGGCTGTCC	2217
Qy	297	AspGluValProLeuLeu-----	-----	302
Db	2218	CCACACTGCTAGCATTTGCCGATCAAGTAG	CGGCAACATAGAATCGGCAATACAATC	2277
Qy	303	-----IleGlyAspValPheHisGly---Lys	311	
Db	2278	CAAGACAAAGACAAATCCAACGCTGCCAG	CATTAAATGATATATTAATACAGGCTTTAAC	2337
Qy	312	TyrGluThrLysLysAsnLeuIleGluAsnAla	SerAlaGluHisGlyTyrPheAspGly	331
Db	2338	CTAAAAAATAAACAACCCCATTTGCTCTCC	-----	2373

QY	332	ArgTrpLeuAspArgSerValAspValle-----LeuProAspAsnThrAlaAspVal	349
DB	2374	-----ACTTATGACATTGTTGACTTTGCCAAATGCCAATGCCACCACCGCC	2418
QY	350	SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe---PheThr	368
DB	2419	ACAGTACCAACCATGATACCGCTAC-----AAACACAGTAAAGTGGTATATGATGTGAAT	2472
QY	369	IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys	386
DB	2473	GTGGATGATACAAACCATTCATCTAAACAGGCACCTGATCACAAATAAAAAAATTCGCGTCMA	2532
QY	387	ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla	406
DB	2533	ACCACCAAACTGAAC-----	2547
QY	407	ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr	425
DB	2548	---AAAAACAAGTCTAATGGTAAATACAGCAACTAATCTTAATGTTAACTCTACTGATGAA	2604
QY	426	GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln	444
DB	2605	GATGCCCTTGTAAACGCCAAAGACATCGCGGAAATCTAAACACCCCTAGCCAAGGAATTT	2664
QY	445	SerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProGluIle	464
DB	2665	CACACCACCAAGGCACACGACAGCACCGCCCTACAACACCTTTACCGTTAAAAAGGTAGAT	2724
QY	465	GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla	484
DB	2725	GAATAATAATTAATGCTGATGACGCCCAACGCCCATCACCGTGGT-----CAAAAGAAGCGCA	2778
QY	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaIleAlaLysAlaArgHisLeuTyrAsp	504
DB	2779	ATAAT-----CAAGTCAACACCCCTAACACTCAAAGGTGAACCGCTCTTAAT	2826
QY	505	MetProAspAspArg-----ValLeuAlaIleAsnHisAspAspGlyValAsn	520
DB	2827	ATTAAACCGCAGCAAAATGGTACGGTTAGCTTTGGCATTAACACCACCAAGCGCTCTTAA	2886
QY	521	ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro	540
DB	2887	-----GCCGGCAAAAGCACCCCTAAAC	2907
QY	541	AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys	560
DB	2908	GACGGTGGC-----TTGTCTATTAACCAACCCACTGGTAGCGCAACAAATCCAA	2955
QY	561	ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp	580
DB	2956	GTCGGTGTGATGGCTGAAGTTGCCAAGGTT-----	2988
QY	581	LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal	600
DB	2989	---AATAATAATAGTGTGTAGTGCTGGC-----ATTGATGGCAGCAACTCGCATTT	3036
QY	601	ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg	620
DB	3037	ACCAAGATGAATGGCTTTACTGGGACTAATGGC-----	3072
QY	621	LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu	640
DB	3073	---TCACCTTGATAAA-----AGCAAAACC-----	3093
QY	641	AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis---SerThr	659
DB	3094	-----CACCTAAGCAAA	3105
QY	660	AsnGlyPheAspLeuSerThrThrLeuGluHisGluIleSerArgSerIleIleGln	679
DB	3106	GACGGCATTAACGGCAGGTGTTAAAAAGATTAACCAACATTCATCAATCGGTGATGTGCCCAA	3165
QY	680	Asn-----GlyGlyTrpAsnArgThrTyrSerLeuLeuArgTyrArgLeu	693

Db 3166 AACAGCCATGATGCTGTACAGCGGC-----AAGATTTATGATTTAAACCGCAACTT 3219  
Qy 694 AspLysLeuLysThrGlnAlaProProGluThrTpGlnAspLeuProValAspPheVal 713  
Db 3220 GAA----- 3222  
Qy 714 AsnGlyLysProSerGlnAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal 732  
Db 3223 ---AACAAATCAGCAGTACTGCCAAACAGACAACTCATTACAGAAATTCAGTA 3279  
Qy 733 AlaAsp----- 3279  
Db 3280 GCAGATGAACAAAGTAACTTTACGGTTAGTAACCCCTTACTCCAGTTATGACACCTCA 3339  
Qy 747 TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet 762  
Db 3340 AAGACCTCTGATGTCATCACCTTCGAGTGAAACGCGCATTTACCAAGGTAATAATA 3399  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 3400 GGTGTGGTGGCGTGTGGCATT-----GACCAAAACCAAGGCTTA 3438  
Qy 783 AsnArgAlaHisGlnMetThrClyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPhe 802  
Db 3439 ACCAGCCCTAAGCTGACCGTGGGTAAATAATGCAAGGCAATGTGTCATGACAGCCAA 3498  
Qy 803 AsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyr 822  
Db 3499 AAT-----GGTCAAAATACCATCACAGGACTA 3525  
Qy 823 AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842  
Db 3526 AGCAACACTCTAGTAAATGTTACCAATGATAAGCT----- 3561  
Qy 843 AlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp-----LeuArgLeu 859  
Db 3562 AGCGTACGCACACAGACAGAGCGC---AATAATATCAAGAGAGACAAACCGCGTCC 3618  
Qy 860 AlaValPheGlyAspIle-----GlyAsnAlaTyr 869  
Db 3619 GCCACCATGTTGTATGCTGAAGCCAGCGCTTAACCTGCAAGGCAATGGTGAACGGTT 3678  
Qy 870 AspLysGlyPheThrAsnAspThr 877  
Db 3679 GACTTTGTCTCCACTTATGACACC 3702  
RESULT 25  
US-08-478-370-1  
; Sequence 1, Application US/08478370  
; Patent No. 5808024  
; GENERAL INFORMATION:  
; APPLICANT: SASAKI, Ken  
; APPLICANT: HARKNESS, Robin E.  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER  
; MEMBRANE PROTEIN OF MORAXELLA  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,370

FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-502  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6973 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-478-370-1  
Alignment Scores:  
Pred. No.: 0.000858 Length: 6973  
Score: 136.00 Matches: 191  
Percent Similarity: 32.95% Conservative: 128  
Best Local Similarity: 19.73% Mismatches: 351  
Query Match: 2.88% Indels: 298  
DB: 1 Gaps: 49  
US-09-914-168-2 (1-919) x US-08-478-370-1 (1-6973)  
Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHis 41  
Db 2064 CCATATTGGTAAATAAACAACCTTAAGTGGGTAGTGTGGCAATTACCATAGACATGGC 2123  
Qy 42 ValProAlaHisAspThrAlaIle---AsnGlnAlaLysAlaGlyAsnProProValLeu 60  
Db 2124 ATTGATGCGAGGTAATAAAGATCAGTAATCTTGCCAAAGGTAGCAGTCTAACGATGCG 2183  
Qy 61 LeuThrProGluGlnIleGlnAla-----ArgLeuAsnAla---AlaGlyLeuAsn 76  
Db 2184 GTTACATCGAAGACAGCTCAAGCGCTACTTTAAAGCGAGCGCTGGCATCAGT 2243  
Qy 77 AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspAspGlnSerProIleSer 96  
Db 2244 GTCACACCTACTGAATATATCAGTTGATGCTAAGAGTGGCAATGTACCGCCCACTTAC 2303  
Qy 97 ArgIleGlyGluGlnSerProProLeuGlyLeuAsp-----MetSerVal 111  
Db 2304 AACATTGGCGTGAACACCAACCGAGCTTAACAGTGGCACTAGTGATAAATTTAGTGT 2363  
Qy 112 IleGluGluThrThrProLeuSerLeu-----GluGluLeuPheAlaGlnGluSerThr 129  
Db 2364 AAGGTAGTGGTACGAACAACATAGCTTAGTTACCGCGCAACATTTGGCAAGCTATCTAAAT 2423  
Qy 130 GluMetGlyIleAsnProAsnAsp-----TyrIleProGluTyrGlnGly 144  
Db 2424 GAAGTCAATCGAAGCGCTGACAGTCTCTACAAAGCTTTACCGTTAAAGCAAGACGAT 2483  
Qy 145 GluGlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeu 164  
Db 2484 GATGACGCCACAGCTATCACCGTGGCTAAAGATACACAAAA-----J----- 2525  
Qy 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys 184  
Db 2526 -----AATGCCGCGCAGTCAGCATCTTAAACACTCAAA 2558  
Qy 185 AlaLys-----PheTyrGlnSer 190  
Db 2559 GGTAAACACGGTCTAACGGTTGCTACCAAAAAAGATGGTACGGTTACCTTTGGCGCTAGC 2618  
Qy 191 SerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla 210  
Db 2619 CAAGATAGCGTCTGACC-----ATTGCAAAAGCACCCCTAAACACGATGCG---TTG 2669  
Qy 211 AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla-----Met 225





```
Qy 870 AspLysGlyPheThrAsnAspThr 877
   |||
Db 4386 GACTTTGTCTCCACTTATGACACC 4409
RESULT 26
US-08-483-855C-1
; Sequence 1, Application US/08483855C
; Patent NO. 6440424
; GENERAL INFORMATION:
; APPLICANT: Sasaki, Ken
; APPLICANT: Harkness, Robin E.
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; FILE REFERENCE: 1038-503 MIS
; CURRENT APPLICATION NUMBER: US/08/483,855C
; CURRENT FILING DATE: 1995-06-07
; PRIOR FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6973
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-08-483-855C-1
Alignment Scores:
Pred. No.: 0.000858 Length: 6973
Score: 136.00 Matches: 191
Percent Similarity: 32.95% Conservative: 128
Best Local Similarity: 19.73% Mismatches: 351
Query Match: 2.88% Indels: 298
DB: 4 Gaps: 49
US-09-914-168-2 (1-919) x US-08-483-855C-1 (1-6973)
Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleLeuHis 41
   |||
Db 2064 CCATATTGGATAAAACAACTTAAAGTGGTAGTGTGCAATTAACCATAGACAATGCC 2123
Qy 42 ValProAlaHisAspThrAlaIle---AsnGlnAlaLysAlaGlyAsnProValLeu 60
   |||
Db 2124 ATTGATGCGAGTAATAAAGATCAGTAATCTGCCAAAGTAGCAGTGCTAACGATCGC 2183
Qy 61 LeuThrProGluGlnIleGlnAla-----ArgLeuAsnAla---AlaGlyLeuAsn 76
   |||
Db 2184 GTTACCATCGAACAGCTCAAGCCGCCAAGCCTACTTTAAACGCGAGCGCTGGCATCAGT 2243
Qy 77 AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspAspGlnSerProIleSer 96
   |||
Db 2244 GTCACACCTACTGAAATATACAGTTGATGCTAAGAGTGGCAATGTTACCGCCCAACTTAC 2303
Qy 97 ArgIleGlyGluGlnSerProLeuGlyLeuAsp-----MetSerVal 111
   |||
Db 2304 AACATTGGCGTGAACACCCAGCCTTAACAGTATGATGCAATTAATTAAGTTAGTGT 2363
Qy 112 IleGluGluThrThrProLeuSerLeu-----GluGluLeuPheAlaGlnGluSerThr 129
   |||
Db 2364 AAGGCTAGTGTAGCAACAATAGCTTAGTTACCGCCGCAACATTTGGCAAGCTATCTAAAT 2423
Qy 130 GluMetGlyIleAsnProAsnAsp-----TyrIleProGluTyrGlnGly 144
   |||
Db 2424 GAAGTCAATCGAACCGGTGACAGTGTCTACAAAGCTTTTACCGTCTAAAGAGAAGACGAT 2483
Qy 145 GluGlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeu 164
   |||
Db 2484 GATGACGCCAACGCTATACGCTGGCTTAAAGATACGACAAAA-----2525
Qy 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys 184
   |||
Db 2526 -----ANTCGCGCGCAGTCAGCATCTTTAAACTCAA 2558
Qy 185 AlaLys-----PheTyrGlnSer 190
   |||
Db 2559 GGTAACAAACGGTCTAACGGTTCTTACCAAAAGATGTGACGGTTACCTTTGGGCTTAGC 2618
Qy 191 SerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla 210
   |||
Db 2619 CAAGATAGCGTCTGACCC-----ATTGCAAAAGACCCCTAAACAACGATGCC---TTG 2669
Qy 211 AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla-----Met 225
   |||
Db 2670 ACTGTTAAAGATACCAACGACAAATC---CAAGTCGGTGTCTAATGGCATTAATTTACT 2726
Qy 226 AspLeuAsnGlySerIlePro-----ArgLeuArgGlnThrAlaLeuValAlaAlaArg 243
   |||
Db 2727 AATGTGAATGTAGTAATCCAGGTACTGCGATTCGCAATACCGCTCGCATTAACAGAGAT 2786
Qy 244 AlaValGlyTyrTyrAspLeuSerIleLeuArgAsnSerIleGlyGluValAsp 263
   |||
Db 2787 AAAATTGGCTTTGGTGGTCTGATGGTGCAGTT-----2819
Qy 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal---282
   |||
Db 2820 -----GATACAAACAACT---TATCTTGATCAAGACAGCTACAAGTTGGC 2864
Qy 283 -----ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296
   |||
Db 2865 AATGTTAAGATTACCAACACTGGCATTAACGCGAGTGGTAAAGCCATCACAGGCTGTCC 2924
Qy 297 AspGluValProLeuLeu-----302
   |||
Db 2925 CCAACACTGCCTAGCATTTGCCGATCAAGTAGCCCAACATAGAACTGGGCAATACAATC 2984
Qy 303 -----IleGlyAspValPheHisGly---Lys 311
   |||
Db 2985 CAAGCAAAAGCAAAATCCACGCTGCCAGCATTAATGATATATTAATACAGCGCTTTAAC 3044
Qy 312 TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331
   |||
Db 3045 CTAAATAAATAAACAACGCCATTCGATGCTGCTCC-----3080
Qy 332 ArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsnThrAlaAspVal 349
   |||
Db 3081 -----ACTATGACATTTGTGCTTGGCAATGGCAATGCCACCGCCCGCC 3125
Qy 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe---PheThr 368
   |||
Db 3126 ACAGTAACCCATGATACCGCTAAC-----AAACACGATTAAGTGGTATATGATGTGAAT 3179
Qy 369 IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys 386
   |||
Db 3180 GTGGATGATACAAACCATTCATCTAACAGGCATGTGACAAATAAAAACTTGGCGTCAAA 3239
Qy 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406
   |||
Db 3240 ACCACCAACTGAC-----3254
Qy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr 425
   |||
Db 3255 ---AAAAAGTGTCTAATGGTAATACAGCACTACTTTAATGTTAACTCTAGTATGATGAA 3311
Qy 426 GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln 444
   |||
Db 3312 GATGCCCTTGTAAACGCCCAAGACATCGCGAAAAATCTAACACCCCTACCAAGGAATTT 3371
Qy 445 SerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464
   |||
Db 3372 CACACCAACCAAGGCACAGACACCGCCCTTACAAACCTTTTACCGTTTAAAGGTAGAT 3431
Qy 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484
   |||
Db 3432 GAAATAATAATGCTGATGACGCCCAACCGCATCCCGTGGGT-----CAAAGAACGCA 3485
```

Qy	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp	504
Db	3486	ATAATA-----CAAGTCAACACCCCTAAACAGGTGAACACGGTCTTTAAT	3533
Qy	505	MetProAspArg-----ValLeuAlaIleAsnHisAspAspGlyValAsn	520
Db	3534	ATTAAAAACCCACAAAAAATGGTCAGGTACCTTTGGCATTTAAACACCAACGCGTCTTAAA	3593
Qy	521	ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro	540
Db	3594	-----GCGCGCAAAAGCACCTTAAC	3614
Qy	541	AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys	560
Db	3615	GACGGTGGC-----TTGTCTATTAAAAACCCCACTGGTAGCGAACAAATCCAA	3662
Qy	561	ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp	580
Db	3663	GTCGGTCTCATCGCTGAAGTTCGCCAAGTT-----	3695
Qy	581	LysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuVal	600
Db	3696	---AATAATAATGGCTGTGTAGTGCTGGC-----ATTGATGGCACAACTCGCATT	3743
Qy	601	ThrLysPheGluHisAsnLeuLeuAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg	620
Db	3744	ACCAGAGATGAATTCGCTTTACTGGGACTAATGCG-----	3779
Qy	621	LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu	640
Db	3780	---TCACCTTGATAAA-----ACCAAAACCC-----	3800
Qy	641	AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis---SerThr	659
Db	3801	-----CACCTAAGCAAA	3812
Qy	660	AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGln	679
Db	3813	GACGGCATTAACGCAGGTGTAAAGAGATTACCAACATTCATCAGGTGAGATTGCCCAA	3872
Qy	680	Asn-----GlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeu	693
Db	3873	AACAGCCATGATCCTGTGACAGCGCGC-----AAGATTTATGATTTAAAAACCGAACTT	3926
Qy	694	AspLysLeuLysThrGlnAlaProProGluThrTyrGlnAspLeuProValAspPheVal	713
Db	3927	GAA-----	3929
Qy	714	AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal	732
Db	3930	---AACAAAAATCAGCAGTACTGCCAAAACAGCACAAAACCTCATTTACAGAAATCTCAGTA	3986
Qy	733	AlaAsp-----AsnLeuValAsnProMetArgGlyTyrArgGlnArg	746
Db	3987	GCAGATGAACAAGGTAACTTAACGTTAGTAGTAAACCCCTTACTCCAGTTATGACACCTCA	4046
Qy	747	TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet	762
Db	4047	AAGACCTCTGATCTCATCACCCTTCAGGTGNAACGGCATTTACCAACAGGTAAATAAAA	4106
Qy	763	AlaIleAlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer	782
Db	4107	GGTGTGGTGGCTCTGGGCATT-----GACCAAAACCAAGGCTTA	4145
Qy	783	AsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrPheSerAspAsnPhe	802
Db	4146	ACCAGCCCTAAGCTGACCGTGGGTAAATAATGCAAGGACCATTTGTCATGACAGCCAA	4205
Qy	803	AsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyr	822
Db	4206	AAT-----GGTCAAAATACCATCACACGACTA	4232
Qy	823	AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnValLeu	842

[illegible]

```

RESULT 27
US-08-621-944A-1
: Sequence 1, Application US/08621944A
: Patent No. 6440425
: GENERAL INFORMATION:
: APPLICANT: SASAKI, Ken
: APPLICANT: HARKNESS, Robin E.
: APPLICANT: LOOSMORE, Sheena M.
: APPLICANT: KLEIN, Michel H.
: TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER
: TITLE OF INVENTION: MEMBRANE PROTEIN OF MORAXELLA
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & Mcburney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/621,944A
: FILING DATE: 26-MAR-1996
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/478,370
: FILING DATE: 07-JUN-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-587
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6973 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 708..6683
: US-08-621-944A-1

```

Alignment Scores:		
Pred. No.:	0.000858	Length:
Score:	136.00	Matches:
Percent Similarity:	32.9%	Conservative:
Best Local Similarity:	19.73%	Mismatches:
Query Match:	2.88%	Indels:
DB:	4	Gaps:
		6973

[illegible]



Db 2727 AATGTGAATGGTAGTAATCCAGGTACTGGCATTGCAAAATACCGCTCGCATTAACAGAGAT 2786  
Qy 244 AlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAsp 263  
Db 2787 AAAATGGCTTGGCTGGTCTGATGGTGAGTT----- 2819  
Qy 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal--- 282  
Db 2820 -----GATACAAACAAACCT---TATCTTGATCAAGACAAAGCTACAAGTTGGC 2864  
Qy 283 -----ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296  
Db 2865 AATGTTAGATTACCAACACTGGCATTACGAGGTGGTAAAGCCATCACAGGGTGTC 2924  
Qy 297 AspGluValProLeuLeu----- 302  
Db 2925 CCAACACTGCCTAGCATTCGCCGATCAAGTACGCCCAACATAGAACTGGGCAATACATC 2984  
Qy 303 -----IleGlyAspValPheHisGly---Lys 311  
Db 2985 CAAGACAAAGACAAATCCAAACGCTGCCAGCATTAATGATATATAATACAGGCTTTAAC 3044  
Qy 312 TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331  
Db 3045 CTAATAAATAATACAAACCCCACTGACTTGTCTCC----- 3080  
Qy 332 ArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsnThrAlaAspVal 349  
Db 3081 -----ACTATGACATTGTTGACTTGGCAATGGCAATGCCACCACCGCC 3125  
Qy 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValPhe---PheThr 368  
Db 3126 ACAGTAACCCATGATACCGCTAAC-----AAAACAGTAAAGTGTGATATGATGTAAT 3179  
Qy 369 IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys 386  
Db 3180 GTGGATGATACAAACCTTCATCATACAGGCACTGATGACAATAAANAACCTTGGCGTCAA 3239  
Qy 387 ArgGluLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406  
Db 3240 ACCACCAAACTGAAC----- 3254  
Qy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr 425  
Db 3255 -----AAAACAAAGTGCTAATGGTAATACAGCACTAACTTAATGTTAACTCTAGTGTAA 3311  
Qy 426 GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln 444  
Db 3312 GATGCCCTTGTAAAGCCAAAGACATCGCCGAAATCTAAACACCCCTAGCCAAAGAAAT 3371  
Qy 445 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464  
Db 3372 CACACACCAAGGACAGACAGACCGCCCTACAAACCTTTACCCTTTAAAGGTAGAT 3431  
Qy 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484  
Db 3432 GAAATAATAATAGCTGATGACGCCCAACGCCATCATCCGTGGGT-----CAAAAGAACGCA 3485  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 3486 AATAAT-----ValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Qy 505 MetProAspAspArg-----ValLeuAlaIleAsnHisAspAspGlyValAsn 520  
Db 3534 ATTAACCGACAAAATGGTAGCTTACCTTGGCATTAAACACCAAGCGGTCTTAA 3593  
Qy 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540  
Db 3594 -----GCCGCAAAAGACACCCCTAAAC 3614  
Qy 541 AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560  
Db 3594 -----GCCGCAAAAGACACCCCTAAAC 3614

Db 3615 GACGGTGGC-----TTGTCTATTAAAAACCCCACTGGTAGCGAAACAAATCCAA 3662  
Qy 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580  
Db 3663 GTCGGTGTGATGGCGTGAAGTTTGCCAAGTT----- 3695  
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600  
Db 3696 ---AATAATAATAGTGTGTAGTGTCTGGC-----ATTGATGGCACAACTCGCAT 3743  
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620  
Db 3744 ACCAGAGATGAATGGCTTACTGGGACTAATGGC----- 3779  
Qy 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640  
Db 3780 ---TCACTTGATAAA-----AGCAAAACCC----- 3800  
Qy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis---SerThr 659  
Db 3801 -----CACCTAAGCAAA 3812  
Qy 660 AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGln 679  
Db 3813 GACGGCATTAACGAGGTGGTAAAGATTACCACTTCAATCAGGTGAGATTGCCCAA 3872  
Qy 680 Asn-----GlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeu 693  
Db 3873 AACGCCATGATGCTGTGACAGCGCG-----AAGATTATGATTTAAACCGCACTT 3926  
Qy 694 AspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheVal 713  
Db 3927 GAA----- 3929  
Qy 714 AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal 732  
Db 3930 ---AACAAATCAGCAGTAGTACTGCCAAAACAGCACAAACTCATTCACCAAGTTCTCAGTA 3986  
Qy 733 AlaAsp-----AsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 3987 GCAGATGAACAAAGTAACTTACGGTTAGTAACCCCTTACTCCAGTTATGACACCTCA 4046  
Qy 747 TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet 762  
Db 4047 AAGACCTCTGATGTCTATCACCCTTGCAGGTGAAACGGCATTACCACCAAGGTAAATAA 4106  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 4107 GGTGTGGTCCGTGTGGCATT-----GACCAAAACCAAGGCTTA 4145  
Qy 783 AsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPhe 802  
Db 4146 ACCACGCCCTAAGCTGACCGTGGTGAATAATAATGCGAAAGGCATTGTCTATTGACAGCAA 4205  
Qy 803 AsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyr 822  
Db 4206 AAT-----GGTCAAAATACCATCACAGGACTA 4232  
Qy 823 AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842  
Db 4233 ACACACACTCTAGCTAATGTTACCAATGATAAGGT----- 4268  
Qy 843 AlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp-----LeuArgLeu 859  
Db 4269 AGCGTAGCCACCAAGACAGAGCGGC---AATAATAATCAAGACGACGACAAACCCGTGCC 4325  
Qy 860 AlaValPheGlyAspIle-----GlyAsnAlaTyr 869  
Db 4326 GCACGATTTGTGTGTGCTGAAGCGCAGGCTTTAACTTCAAGGCAATGTGTAAGCGGTT 4385  
Qy 870 AspLysGlyPheThrAsnAspThr 877  
Db 4386 GACTTGTCTCCACTTATGACACC 4409

## RESULT 29

US-08-431-718C-1  
; Sequence 1, Application US/08431718C  
; Patent No. 6335018  
; GENERAL INFORMATION:  
; APPLICANT: Ken, Sasaki  
; APPLICANT: Robin, Harkness E.  
; APPLICANT: Loosmore, Sheena M.  
; APPLICANT: Klein, Michel H.  
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF  
; FILE REFERENCE: 1038-429 MIS  
; CURRENT APPLICATION NUMBER: US/08/431.718C  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 6975  
; TYPE: DNA  
; ORGANISM: Moraxella catarrhalis  
US-08-431-718C-1

Alignment Scores:  
Pred. No.: 0.000859 Length: 6975  
Score: 136.00 Matches: 191  
Percent Similarity: 32.95% Conservative: 128  
Best Local Similarity: 19.73% Mismatches: 351  
Query Match: 2.88% Indels: 298  
DB: 4 Gaps: 49

US-09-914-168-2 (1-919) x US-08-431-718C-1 (1-6975)

QY 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHis 41  
DB 2064 CCATATTGGGATAAAAACAACTTAAAGTGGGTAGTGTGCAATTACCATAGACAATGGC 2123

QY 42 ValProAlaHisAspThrAlaIle-----AsnGlnAlaLysAlaGlyAsnProProValLeu 60  
DB 2124 ATTGATCGAGGTATAAAAGATCAGTAATCTTCCAAAGGTACAGTGTAAAGATGGC 2183

QY 61 LeuThrProGluGlnIleGlnAla-----ArgLeuAsnAla-----AlaGlyLeuAsn 76  
DB 2184 GTTACCATCGAACACAGCTCAAGCGCCAGCCTACTTTAAACGCGAGCGTGGCATCAGT 2243

QY 77 AlalysProGlnSerGlnAlaLeuAspValValIasnPheAspAspGlnSerProIleSer 96  
DB 2244 GTCACACCTACTGAATATCAGTTGTGCTAAGAGTGGCAATGTATCCGCCCAACTTAC 2303

QY 97 ArgIleGlyGluGlnSerProProLeuGlyLeuAsp-----MetSerVal 111  
DB 2304 AACATTGGCGTGAACACCCAGCTTAACAGTGTACAGTGTACGATGATAATTTAGTGT 2363

QY 112 IleGluGluThrThrProLeuSerLeu-----GluGluLeuPheAlaGlnGlnSerThr 129  
DB 2364 AAGGTAGTGGTACGAACAATAGCTTAGTTACCGCGCAACATTTGGCAAGCTATCTAAAT 2423

QY 130 GluMetGlyIleAsnProAsnAsp-----TyrIleProGluTyrGlnGly 144  
DB 2424 GAAGTCAATCGAAGCGGTGACAGTGTCTACAAAGCTTTACCGTTAAAGAAAGAGAGAT 2483

QY 145 GluGlnProAsnSerGluValValValProProThrLeuGluProGluLysProGlyLeu 164  
DB 2484 GATGACGCCCAACGCTATCACCCTGGTGGTAAAGATACGACAAA----- 2525

QY 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys 184  
DB 2526 -----AATGCGCGCGATCGATCGATCTTAAACACTCAA 2558

QY 185 AlaLys-----PheTyrGlnSer 190  
DB 2559 GGTAAAAAGGCTAACGGTTGCTACCAAAAAAGATGGTACGGTTACCTTTGGCTTACG 2618

QY 191 SerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla 210  
DB 2619 CAAGATACGGTCTGAGC-----ATTGGCAAAAGCACCCCTAAACACAGCATGGC---TTG 2669

QY 211 AsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAla-----Met 225  
DB 2670 ACTGTTAAAGATACCAACGACAAATC---CAAGTCGGTGTCTAATGGCATTTAAATTTACT 2726

QY 226 AspLeuAsnGlySerIlePro-----ArgLeuArgGlnThrAlaLeuValAlaAlaArg 243  
DB 2727 AATGTGAATGATAGTAATPCCAGTACTGGCATTTGCAATACCCCTGCGCATTTCCAGAGAT 2786

QY 244 AlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAsp 263  
DB 2787 AAAATTGGCTTTGCTGTTCTGTGTCAGTT----- 2819

QY 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal--- 282  
DB 2820 -----GATACAAACAAACCT---TATCTTGATCAAGACAAGCTACAGTTGGC 2864

QY 283 -----ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296  
DB 2865 AATGTTAAGATTACCAACACTGGCATTAACGCGGTGTTAAAGCCATCACAGGCGCTGTC 2924

QY 297 AspGluValProLeuLeu----- 302  
DB 2925 CCAACACTGCGCTAGCATTTGCGGATCAAAGTAGCGCAACATAGAACTGGGCAATACAACT 2984

QY 303 -----IleGlyAspValPheHisHisGly---Lys 311  
DB 2985 CAGACAAGACAAATCCAACGCTGCCAGCATTAATGATATATATAATACAGGCTTTTAAAC 3044

QY 312 TyrGluThrIlyLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331  
DB 3045 CTAATAAATAATAACAACCCCATGACTTTGTCTCC----- 3080

QY 332 ArgTrpLeuAspArgSerValAspValle-----LeuProAspAsnThrAlaAspVal 349  
DB 3081 -----ACTTATGACATTGTTGACTTTGCCAATGGCAATGCCACACCGCC 3125

QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe---PheThr 368  
DB 3126 ACAGTAACCATCATATACCGCTAAC-----AAACACGATAAGTGGTATATGATGTGAAT 3179

QY 369 IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys 386  
DB 3180 GTGGATGATACAACCATTCATCTAACAGGCACCTGATCACAATAAAAACTTGCGCTCAA 3239

QY 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406  
DB 3240 ACCACCAACTGAAC----- 3254

QY 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr 425  
DB 3255 ---AAAAAAGTGCTAATGGTAATACAGCACTAACTTTAATGTTAACTCTAGTGATGA 3311

QY 426 GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln 444  
DB 3312 GATGCGCTTTGTTAAACGCCAAAGACATCGCGGAAAATCTAAACACCCCTAGCAAGGAAT 3371

QY 445 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464  
DB 3372 CACACCCCAAGGCACACGACGACCGCCCTACAAACCTTTACCGTTTAAAGGTAGAT 3431

QY 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484  
DB 3432 GAAAATAATAATGCTGATCAGCCCAACGCAATCACCGTGGT-----CAAAAGAACGCA 3485

QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
DB 3486 AATAAT-----CAAGTCAACACCCCTAACACATCAAAAGGTCAAAAGCTTCTTAAT 3533

QY 505 MetProAspAspArg-----ValLeuAlaIleAsnHisAspGlyValAsn 520

Db 3534 ATTTAAACCGCAAAAATGGTACGGTTACCTTTGGCATTAACACACACAGCGGCTCTTAA 3593  
Qy 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540  
Db 3594 -----GCGGCAAAAGACCCCTTAAC 3614  
Qy 541 AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560  
Db 3615 GACGGTGGC-----TTGCTATTAAAAACCCACTGGTAGCGCAACAATCCAA 3662  
Qy 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580  
Db 3663 GTCGGTGTGTAGCGGTGAAGTTTGCCAAAGTT----- 3695  
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600  
Db 3696 ---AATAATATAGGTGTGTAGGTCTGGC-----ATTGATGGCACAACTCGCAAT 3743  
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620  
Db 3744 ACCAGAGATGAATTTGGCTTTACTGGGACTAATGCG----- 3779  
Qy 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640  
Db 3780 ---TCACTTGTATAA-----AGCAAAACCC----- 3800  
Qy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHis---SerThr 659  
Db 3801 -----CACCTAAGCAAA 3812  
Qy 660 AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleLeuGln 679  
Db 3813 GACGGCATTAAACGACAGGTGGTAAAGATTAACCAACATTCATACAGGTGAGATTGCCCAA 3872  
Qy 680 Asn-----GlyClyTrpAsnArgThrTyrSerLeuArgTyrArgLeu 693  
Db 3873 AACAGCCATGATGCTGTGCACAGCGGC-----AAGATTTATGATTTAAAAACCGAAGCTT 3926  
Qy 694 AspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheVal 713  
Db 3927 GAA----- 3929  
Qy 714 AsnGlyLysProSerGlnAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal 732  
Db 3930 ---AACAAAAATCAGCAGTACTGCCAAAACAGCACAAAACCTCATTACAGCAATTCAGTA 3986  
Qy 733 AlaAsp-----AsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 3987 GCAGATGAACAAGGTAATAACTTTACGGTTTAGTAACCCCTTACTCCAGTTATGACACCTCA 4046  
Qy 747 TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet 762  
Db 4047 AAGACCTCTGATGTCATCACCTTCAGGTGAAAACGCGCATTTACCACCAAGTAATAAA 4106  
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782  
Db 4107 GGTGTGGTGGGTGGGCATT-----GACCAAAACCAAGGCTTA 4145  
Qy 783 AsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPhe 802  
Db 4146 ACCAGCGCTAAGCTGACCGTGGGTAAATAATGCAAGGCAATTTGCTATTGACAGCAAA 4205  
Qy 803 AsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyr 822  
Db 4206 AAT-----GGTCAAAATACCATCACAGGACTA 4232  
Qy 823 AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842  
Db 4233 AGCAACACTCTAGCTAATGTTACCAATGATAAAGT----- 4268  
Qy 843 AlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAsp-----LeuArgLeu 859  
Db 4268 -----LeuArgLeu 859

Db 4269 AGCGTACCCACACAGAACAGGGC---AATATAATCAAAAGACAGACAAAACCGTGCC 4325  
Qy 860 AlaValPheGlyAspIle-----GlyAsnAlaTyr 869  
Db 4326 GCCAGCATTTGTAGTGTCTAAGCCAGCGCTTAACCTTGCRAAGGCAATGGTGAAGCGGTT 4385  
Qy 870 AspLysGlyPheThrAsnAspThr 877  
Db 4386 GACTTTGTCTCCACTTATGACACC 4409  
RESULT 30  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2  
Alignment Scores:  
Pred. No.: 66.9 Length: 4403765  
Score: 136.00 Matches: 191  
Percent Similarity: 33.98% Conservative: 107  
Best Local Similarity: 21.78% Mismatches: 284  
Query Match: 2.88% Indels: 298  
Gaps: 47  
US-09-914-168-2 (1-919) x US-09-103-840A-2 (1-4403765)  
Qy 8 AlaAsnArgSerPheMetProValAlaLeuAlaTyrLeuProLeuMet----- 24  
Db 462238 CGGTCCCGGTCTCGAGGCCACGACCTCGATTCTAAGCTCGCGCGGAAACGGCGGG 462179  
Qy 25 -----ThrSerGlnAlaLeuAlaGlnGlnAsnPro 35  
Db 462178 ATGCACAAAGTGGTAAATTGAGCGGAACAGACTCAACATTGACGCGTTGAACAACCG 462119  
Qy 36 AlaAsnIleAlaAsnHisValPro----- 43  
Db 462118 ACAAGCATTTCAAGCGG-ACCCCGAATGGAGGTGCTGTGAGTCTGTTAACCCGAGACC 462060  
Qy 44 -----AlaHisAspThrAlaIleAsnGlnAla-----LysAlaGlyAsnPro 58  
Db 462059 AAGACGACGGCGGCTAACCGCGGCTTACAGGCGGCTTCACCGCCGCAATCCCGAG 462000  
Qy 59 Val-----LeuLeuThr----- 62  
Db 461999 ATCCGCGCGGCTCACCTGCTGTGATGCGTGTGCTGACCAAAACAGCGGTATCGCCGACCG 461940  
Qy 63 -----ProGluGlnIleGlnAla----- 68  
Db 461939 CTACTGGAGGTGTCGGTGTTCGAGCCCGCGCACCTCGCGCGGAAACACGAGCGCTGCTC 461880  
Qy 69 ---ArgLeu---AsnAlaAlaGlyLeuAsnAlaLysProGln----- 80  
Db 461879 GACCGTTCGCGGCGGACTGGAGCCAGCGCCGCGAGCTGTCCCGCGAGTCGTTA 461820

QY 81 -----SerGlnAlaLeuAspValValAsnPhaAspAspGlnSerProile 95  
Db 461819 GCGGCGATCACCACCGCGCAGCAGCTG---GCCACCGAGCTGGACGACGAG----- 461772  
QY 96 SerArgIleGlyGluGlnSerProLeuGlyLeu-----AspMetSerValile 112  
Db 461771 ---TAGCTCTCCAGGACAGCTGATGTCGGGCTGGCCACCGGTGACTCCGAGCTGCC 461715  
QY 113 GluGluThrThr-----ProLeuSerLeuGluLeuLeuPheAlaGlnGlu 127  
Db 461714 AAGCTGTTGACCGCGCAGCGCGCTCGCGCAGCGCTCGCGGAGCGCTTCTGCTCAAGGTG 461655  
QY 128 SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro 147  
Db 461654 CGCGCAGCGCCCGGCTCACAGCGCCCGAACCGGAGCGGAGCTATCAGCGGCTGCAGAAG 461595  
QY 148 AsnSer-----GluValValValProProThrLeu----- 157  
Db 461594 TACTCCACCGACCTGACCGCGCGCGCGGAGGCAACTCGACCGCGTCTCGCGCGC 461535  
QY 158 GluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAsp----- 175  
Db 461534 GACAACGAGATCCCGCGCTGGTGCAGGTGCTGCTCCGCTCGCACCAAGAACACCCCGTG 461475  
QY 176 -----GlyValAsnLys-----ValProArgLeu 183  
Db 461474 CTGATCGGTGAGCGCGCTCGGCAAGACCCGATCGTGGAGGCGCTGGCGCAGTCTCGGC 461415  
QY 184 LysAlaLysPheTyrGlnSerGlnSerGlyGluThr----- 196  
Db 461414 GTGCGCGCGCAGCTGCGCGAGAGCTTGGCGCACAGACCATCTGCTCGCTCGATCTCGGC 461355  
QY 197 SerAlaIleGlySerSerHisGlnLysThrCluProTyrAlaAsnIleLysAlaLeu 216  
Db 461354 TCGATGTCGCGGCTCCAAATACCGCGGGAATTCGAGGAACGGCTCAAGCGCTCCTC 461295  
QY 217 GluAspIle----- 219  
Db 461294 GACGACATCAAGAACTCGCGCGGCAAAATCATCATCGTTCATCGACGAGCTGCACACCATC 461235  
QY 220 -----ThrGlnGluSerAlaMetAspLeuAsnGlySerIle---ProArg 233  
Db 461234 GTCGGCGCGCGCGCCACCGCGGAGGCGGATGACCGCGCAACATGATCAAGCGCATG 461175  
QY 234 LeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSer 253  
Db 461174 CTGCGCGCGCGCGATTA-----CGGCTGTCGGGCGCACCGCTGGAC----- 461130  
QY 254 IleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProVal 273  
Db 461129 GAATACCGCAAGCACATC---GAGAAGGACCGCGCGCTCGAGCGCGTTCACACAGGTG 461073  
QY 274 TyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThr 293  
Db 461072 TACGTCGGCGAGCGCTCGGTGGAG----- 461049  
QY 294 ThrValAlaAspGluValProLeuLeuIleGly-----AspValPhe-----HisHis 309  
Db 461048 -----CACACCATCGGCATCTTCGCGGGCTCAAGAGCCGCTCAGAGGTGCACAC 460998  
QY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329  
Db 460997 GGGTGGCATCATCCGACTCGCGCTGGTGGCAGCTGCCACTTTGAGCGACCGGTATATC 460938  
QY 330 AspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349  
Db 460937 ACCGCGCGCTTC-----CTGCGCGCAACAGCGCATCGAC--- 460905  
QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspCluValValPhePheThrIle 369  
Db 460904 ---CTGTCGACGAGCGCGCACCGCGCTCGGATGGAG-----ATC 460866

QY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgCluLeu 389  
Db 460865 GACTCGCGCGCGCTCGAGATC---GACGAGTCCGAGCGGCTGGTGGCGGCTCGAGATC 460809  
QY 390 LeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAla 409  
Db 460808 GAAGAGATGCGCTGTCTCCAAAGAAGACGAGCGC---TCGCGGAGCGGCTTCCCAAG 460752  
QY 410 LeuSerAsnAspLeu-----lleAlaThrArgTyrPhe 420  
Db 460751 CTGCGCTCCGAGCTGGCGCGACCAAGAGAGTGGCCGAGCTCACACCGCTGGCAG 460692  
QY 421 AsnMetValAsnThr-----GluIleValPheProGluArgGluGlnIleGlnAsnAspGln 439  
Db 460691 AACGAGAAGAAGCGATCGAAATCGTCCGCGACCTCAAGGAGCAGCTGGAA----- 460641  
QY 440 ValSerPheGluGlnSerSerSerSerSerSerSerSerSerSerSerSerSerSerThr 459  
Db 460640 -----CCCTCGCGCGGGGAATCCGAGCGGCGCAA----- 460611  
QY 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479  
Db 460610 -----CGCGACGCGCACCTGGCCCAAGCGCGCGAG 460581  
QY 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn-----LeuValAlaAla 497  
Db 460580 CTGCGCTACGAGCGCATCCCGCGAGTGGAGAGAAGCTCGACGCGGCTTCCGCCAGGCG 460521  
QY 498 LysAlaArgHisLeuTyrAspMet-----ProAspAsp----- 508  
Db 460520 CAGGCGCGGAGCAGGTGATGCTCAAGGAGGAGTCCGTCGCCACGACATCCCGCGAGGTG 460461  
QY 509 -----ArgValLeu-----Ala 512  
Db 460460 GTGTCGCGTGGACCGCATCCCGCGCGCTGCTGTGAAGCGAGACCGCAAGCTG 460401  
QY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532  
Db 460400 CTGCGCATGAAGACGAGCTGGGCAAGCGGTCATCGGCGAG---AAGCGCGCGGTACC 460344  
QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552  
Db 460343 GCAGTCTCTCATCGGTG-----CGCGCAGCGCGCGCGGCTGTCGAC----- 460299  
QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValPro 572  
Db 460298 -----CCCAACCGCGCCACCGGCGC----- 460278  
QY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGly 592  
Db 460277 ---TTCATGTTCTCGCGCC-GACCGGTGTGGCAAGACCGAGCTGGCCAAAGCGCTGGC 460222  
QY 593 SerAspThrGlyThrArgLeuValThrLys-PheGluHisAsnLeuIleAsnArgAsp-- 611  
Db 460221 -----CGACTTCTCTTCGACGACGAGCGGCGGCTGTCGCGATCCACAT 460177  
QY 611 ----- 611  
Db 460176 GACCGAGTACGCGGAGAGACACACCTGCTCGTGTGATCGCGCGCCCGCGGTATGT 460117  
QY 612 -GlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys-----GlyVally 629  
Db 460116 GGGATACGAGGCGCGGCTGCTACCGAGCGGCTGGCGCGGCTCCMTACACGCTGTG 460057  
QY 629 sLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuG 649  
Db 460056 GCTGTTTCGACGAGATCGAGAAGCGCACCG-----GACGTGTCGAGCTGCTCTGCA 460003  
QY 649 yTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLe 669  
Db 460002 GGTCTCTCGACGAG-----GGCGCGCTCACCGAGCGGCGCGCGCGCTCGACTTCC 459950  
QY 669 uGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArg 685





Db 2119 GTCACCTCTACGCGAAACAGAAAGTCACCAACAATACGCTGGGTCTACGGCGACCGGTACCAATG 2060  
Qy 368 -----ThrIleasp 370  
Db 2059 ACCAGCAATATTCGGGCAACATAACGTCGTTGGTGAGCCGGAAGCGCAAGCTTCGCGAT 2000  
Qy 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys---ArgGluLeu 389  
Db 1999 AATAAAACGTTTGTAGTTATAGTCTACCGGATGAAGTTCGGGCAAGGTAATAGTATA 1940  
Qy 390 LeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGln---AlaValArg 408  
Db 1939 ACCGAGCGCAAAACGATACGGTGGGGCGGAACACATCACGCTACGGATCTCTCCAG 1880  
Qy 409 AlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 1879 GACGCGTTTACAAATGTAATCGCGGTCAACGCGTCAGATTAAAGTGGCAGCGCAACAAC 1820  
Qy 419 -----TyrPheAsnMet 422  
Db 1819 AACATTACGATAGCGGATACGGCTTACACCGATAATAACGGTTATGCGGTACGTTAACCTT 1760  
Qy 423 ValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPhe 442  
Db 1759 CTCAGCACCCAA-----CTGGGGTTTATCAGGTG-----ACGCAACGCTG 1718  
Qy 443 GluGlnSerSerSerArgThrGlu 451  
Db 1717 GACATAACAGTAGTAGTAAGTTGACGTGAATGTGGCAATGGCAAACTCGAGTTAAACA 1658  
Qy 451 ----- 451  
Db 1657 TCATCGAAACCAAGAACTACGGTCCATAATAGTAGGGGTATTACGGTGCACCGCAACGGCG 1598  
Qy 452 -----ProAlaGlnValAspGluSerThrLeuGluProVal 463  
Db 1597 AGAATGCGGGGGTGAATGATGCCGCGCAAAATATCACCTTTAGCGTAACGCGCTGAA 1538  
Qy 464 IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSer 483  
Db 1537 GGTGCAACGCTAGCAATACAGGGAAGTCTTACTGACCACTCA-----GGTCAGGCC 1484  
Qy 484 AlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyr 503  
Db 1483 AAAGTGACGCTGACCACTGACAAAGTGAATGTCTATACCGTTACGGCCATAATGGCCAA 1424  
Qy 504 AspMetPro-----AspAspArgValLeuAlaIleAsnHisAsp 516  
Db 1423 GATGTTCCGTTGAGCCAGGTAAACGGTTGCGGTTAAGCGCAGATGCTAAACCGGCACAT 1364  
Qy 517 -----AspGlyValAsnArgSer--- 522  
Db 1363 GTTGTGAGCGTCTGCTCTCTGACACCATCACCGCCGCGCATCGATAGCAGACC 1304  
Qy 523 IleLeuGlyArgIleSerAsp-----AlaValSerAlaValAlaArgAlaIle 538  
Db 1303 ATCATCTTCAGAGTGAAGAATGATTACGGATTCCCGGTTCCCGGTTCAAGGTGTCGATATTAGTCAT 1244  
Qy 539 LeuProAspGluSerGluAsnGluValIleAspLeuPro----- 551  
Db 1243 GGCTTACACCAAGCGACCGCGGTAGTTAAATATTCCTCAACTACGCGTACCGATCAGTC 1184  
Qy 552 -----GluArgThrAlaLeuAlaAsnArgLys 560  
Db 1183 GGGCAAGTCACGGGACAAATAACCAAGTACATTGGCAGAAACCTTAAACAGTCAATGTGCAA 1124  
Qy 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580  
Db 1123 GTTCTGGCACAGCCAAACCAATCCGCAACCAATACATTGTTGGTCCGCGCAGCGCGATGAA 1064  
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600

Db 1063 -----AGTAAGTCAATT 1052  
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlnLeuArg 620  
Db 1051 TTGAATCCGATGTTGACACTCTGAAGCTGAC---TACCAGCAGCGCAAACTTACG 995  
Qy 621 LeuSer---GluAspLysLysGlyValLysLeuTyrAlaThrLysProLeu----- 636  
Db 994 CTACATTGCAAGACAAGTACGGTAACCGGATAGTACGCTCTGATCATCTGCAATTTGTC 935  
Qy 637 ---SerHisProLeuAsnAspGlnLeuArg---AlaThrLeuGlyTyrGlnGlnGluVal 654  
Db 934 CAGTCAGGCCCTTCGTGAACCTTCTCAAGTTGACGATATTGATTACAGCCCAAGAAAT 875  
Qy 655 PheGlyHisSerThr----- 659  
Db 874 TATGCGGAGTACACCGTCTACTGCGGCAAAAGAGGAACAGCGACACTCATTTCCC 815  
Qy 660 -----AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIle 677  
Db 814 ATGCTGAACGGGTTTCATCAGGCAAACTTAAAGCATATCGCTGAATCTCATCAATCGATA 755  
Qy 678 IleGlnAsnGlyGlyTrp-----AsnArgThrTyrSerLeuArgTyrArgLeu 693  
Db 754 AAAGAAATGTCCCGTCATCTACTGCAAAACCACTACCTTCTCCACGGCT--- 704  
Qy 694 AspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheVal 713  
Db 703 ---AAATCCCGCAGCGAAGCTTTCGAGGAGCGTATTACACACTC----- 602  
Qy 714 AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAla 733  
Db 661 -----AACAAATGATAACTTTGAAGCGGT-----AAAACCGTTGAT 626  
Qy 734 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753  
Db 625 GATTATATGTTTCAAGTTTCACAGTTCACAGGGTTGG-----GTGCTGTCGATGCT 581  
Qy 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773  
Db 580 TCGGGTAAAGTTTCTTTCGCAAAATATC----- 554  
Qy 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791  
Db 553 -----GGCGATCAACAGCTCAGTCAACAATAAGCGCTTCCCGCAGCAAGAGGTACAAAC 500  
Qy 792 -----IleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyr 807  
Db 499 TACCAGACCTTAATTAAGCTGAAGGCTGGTGGTGAATAATGGAAATCATACCAATATC 440  
Qy 808 ArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu 827  
Db 439 TGGCTA-----GCTGCCAATGCGCTC 419  
Qy 828 SerProIleSerAspLysGlyTyr-----LeuThrGlyGlyGln 840  
Db 418 TGTCATGCTAAAAATGATGATATACTTCCTCGCATCATCACACATTTGACGCTCGGCCAA 359  
Qy 841 ValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860  
Db 358 -----AACAAACGCAAGCGGATCA 338  
Qy 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880  
Db 337 CTGTATGTTGAATGGGGGAACGTT-----GGAGCGTTTTCAGTAATTTCCAAATTTACA 284  
Qy 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900  
Db 283 CCG-----GGAGCTTACTGCAGAGT----- 263  
Qy 901 GlyValLysGluGlyAsnProIleLysLeuHisPhePheIle 915  
Db 262 -----GAATCTGATGATTACAGTCGCGCACTACTATGTC 230

## RESULT 32

US-09-134-001C-322  
; Sequence 322, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 322  
; LENGTH: 30549  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-322

## Alignment Scores:

Pred. No.: 0.0146 Length: 30549  
Score: 135.00 Matches: 166  
Percent Similarity: 32.81% Conservative: 106  
Best Local Similarity: 20.02% Mismatches: 295  
Query Match: 2.86% Indels: 262  
Gaps: 4

US-09-914-168-2 (1-919) x US-09-134-001C-322 (1-30549)

Qy 27 GlnAlaLeuAlaGlnGlnAsnProAlaAsn-----IleIleAsnHisValProAlaHis 45  
Db 8083 CAAGCAGCTGAACGAACATACAAATGCGTTATAAATAATGGAGATCGAAC 8142  
Qy 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGln 65  
Db 8143 TCCCAACAAATTAATGATGCTAAA-----AACACAGTTGAACAGGCACAGAGATTAT 8196  
Qy 66 IleGlnAlaArgLeuAsnAlaLagLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85  
Db 8197 GTTGAAGCTAAAGCAACTAGCTGCTGATAGTACAGTTACAGTTCACAAAGC---GCTTATGAT 8253  
Qy 86 ValValAsnPheAsp-----AspGlnSerPro----- 94  
Db 8254 ACGTTAAATAGAGATGTTTAAACAAATGATAAAAGCCAGCATCTGTAAGCGCTATAAT 8313  
Qy 95 -----IleSerArgIleGlyGlnSerProProLeuGlyLeuAspMetSerValIle 112  
Db 8314 GAAGCCATTTCAAATATTAGAAAAGAAATAGATACAGCTAAAGCGGATGCAAGTAGTACT 8373  
Qy 113 GluGluThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGly 132  
Db 8374 TTGCGAAACACCAATCCCTCCGTTCAACAGTTAGACAGCGCTTTAAATAAATAAATACT 8433  
Qy 133 IleAsnPro-----AsnAspTyrIleProGluTyrGlnGlyGlnProAsnSerGlu 150  
Db 8434 GTTCAACCTAAAGTGAATCAAGCAATTCGTTTACTTCAACCAAGAAATAAATTCAGAA 8493  
Qy 151 ValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAla 170  
Db 8494 CTTGTACAAAGTAAA-----AAACGTTTA----- 8517  
Qy 171 ArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla----- 185  
Db 8518 -----CAAGCGCTGTAATGACATACCTCAACACAGAGGTATGACACAAACA 8568  
Qy 186 ---LysPheTyrGlnSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGln 204  
Db 8569 ATTAATAATTATATGACAAACAAGCTGAAGCTGAAGGAGAGCATTTACATCTGCCAAGA 8628

Qy 205 LysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla 224  
Db 8629 GTGATTGAT-----AATGGGGATGCTACAACTCAAGAATTTACTTCTGAAAAATCT 8679  
Qy 225 MetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAla 244  
Db 8680 -----AAAGTAGAGCAAGCAATGCAA--GCTTTAACTAATGCTAA-- 8718  
Qy 245 ValGlyTyrTyrAspIleAspLeuSerIleIleAtqAsnSerIleGlyGluValAspVal 264  
Db 8719 -----AGTATCTGAGA----- 8730  
Qy 265 IleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly 284  
Db 8730 ----- 8730  
Qy 285 GluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGly 304  
Db 8731 -----GCTGATAAGAATGAGTTACAGACTGCATATAACAAA-----TTAATTGAG 8775  
Qy 305 AspValPheHisGlyLys-----TyrGluThrLysLysAsn 317  
Db 8776 AACGTATCTACCAATGGTAAAAACCGCGAGTATACGTCAATACCAACACCAAGCC 8835  
Qy 318 LeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp----- 335  
Db 8836 AGATACAAAAATCAAATTAATGATGCTAAAAATGAAGCGGAGCGAATTTTAGGTAAATGAT 8895  
Qy 336 -----ArgSerValaspValIleLeuProAsp 344  
Db 8896 AATCCACAAGTATACAAAGTAACTCAAGCATTTGAACAAATCAAAGCTATTCAACCAAAA 8955  
Qy 345 AsnThrAlaAspValSerLeuIleTyrAsp-----ThrGlyThr 357  
Db 8956 TTAACAGAGAGCTATCAACATGCTTCAAAACAAGAAATAATACAGAAATAGTCAATGCT 9015  
Qy 358 GlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThr 377  
Db 9016 AAAAAACAGACTTGAATAATGCAGTAAATATACA---GATCCA---ACACACGCTATGACT 9069  
Qy 378 ThrAsp-----ProAspLysLeuProValLysArgGluLeuGlu 391  
Db 9070 CAAGAAACAATTAATAATTACACCGCTAAAAAGCGAGAGCTCAAAATGAAATACAAAA 9129  
Qy 392 GlnLeuLeuThrValAsnMetGlyGluAla----- 401  
Db 9130 CGAACATGATTTAATAATGGAGATGCTACTGCTCAAGATATTTCTTCTGAAAAATCT 9189  
Qy 402 -----TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu----- 414  
Db 9190 AAGTAGACCAAGTATTACAGCATTTACAAAATGCTAAGAATGACTTAAGAGCTGATAAA 9249  
Qy 415 -----IleAlaThrArgTyrPheAsnMet-----ValAsnThrGluIleValPhe 429  
Db 9250 AGAGAATTACAGACTGCATACAAATACTTATACAAAATGTTAATACCAATCGTAAAAAAA 9309  
Qy 430 ProGlu-----ArgGluGlnIleGlnAsnAspGlnValSer 441  
Db 9310 CCATCTAGTATTCAAAACTATAAGTCTGCAACAGCAAAATATCGAAAAAC----- 9357  
Qy 442 PheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461  
Db 9358 -----CAATAATATACCGCTAAAAAATGAAGCACAATAATGTTCTTTGAAAAATACA---AAC 9408  
Qy 462 ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLysSerProIleGlu 481  
Db 9409 CCTACTGTAATGCGAGTGAAGATGCTTTACCTAAGATAAATGCAATTCACACAGAGTT 9468  
Qy 482 PheSerAlaSerAsnLeuIleGlnAspLys----- 491  
Db 9469 ACAAAAGCTATTAAATATATCTCAAGATAAAGAGAGATAATAGCGCACTTCTTAGACCAAAA 9528  
Qy 491 ----- 491

```
Db 9529 GAAAAATTAGATCAAGCGATTAAATAGTCAACCATCACTAAATGGTATGACTCAAGAATCT 9588
Qy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeu 511
Db 9589 ATTAATAATTACACACAAAGCGTAGAGACGACAAATATAGCAAGTTCCTGCTACACT 9648
Qy 512 AlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly-----ArgIleSer 528
Db 9649 ATTATTAATAATGGGATGGCATCTATTGAACAAATAACAGAAATAAATTCGAGTTGAA 9708
Qy 529 AspAlaValSerAlaValAlaArgAla-----IleLeuProAspGluSer----- 543
Db 9709 GAGCAACTAATGACCTTAACGAACCAAAACACATTTAAACGGCAGATACAACTTCTTTA 9768
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 9769 AAAAGTGAAGTACGGAATTA---AGTAGGAGCGCGACACAAACAAACAAAGCCTAGC 9825
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 9826 AGTGT----- 9831
Qy 584 AspGlyGlnIleGlyLeuGlyTyrPglySerAspThrGlyThrArgLeuValThrLysPhe 603
Db 9831 ----- 9831
Qy 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlu 623
Db 9832 -----AGTGCCTTATAACAATACTATTCTTCGCTACAACTGAAATTCACACAGACTGAA 9885
Qy 624 AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSer----- 637
Db 9886 ANTAGCAATAACTATCATC-----AATAAGCCTATTCTGTTGTAAGAAGTAAT 9939
Qy 638 -----HisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGln 651
Db 9940 AATGATTCGATGAAGTAACCAATGAACCAACGCTTAACAGATACAAATTAACCTTATTA 9999
Qy 652 GlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671
Db 10000 CAACCTTTA-----GCGAATAAAGAAAGCTTAAGAAAGCTCGTAATCGACTTGAAGT 10053
Qy 672 GluIleSerArgSerIleIleGlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyr 691
Db 10054 AAAATTAATGAACCC---GTTCAACAGACGCGTATGACTCAACAACTCTGTTGAGAAATTAT 10110
Qy 692 ArgLeuAspLysLeuLysThrGlnAlaProProGluThrTyrGlnAspLeuProValasp 711
Db 10111 AAGCAAGCTAAATAAAGCTCAAAATGAATCTAGTATTGCACAAACTCT----- 10161
Qy 712 PheValAsnGlyLysProSerGlnGlu 720
Db 10162 ATTAATAATGGTATGATCTGATCAA 10188

RESULT 33
US-09-221-017B-1002/C
; Sequence 1002, Application US/09221017B
; Patent No. 644799
; GENERAL INFORMATION:
; APPLICANT: ROSS, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1002:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...9377
; US-09-221-017B-1002

Alignment Scores:
Pred. No.: 0.00211 Length: 9377
Score: 134.50 Matches: 117
Percent Similarity: 34.00% Conservative: 69
Best Local Similarity: 21.39% Mismatches: 188
Query Match: 2.85% Indels: 173
DB: 4 Gaps: 30

US-09-914-168-2 (1-919) x US-09-221-017B-1002 (1-9377)
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg----- 521
Db 2702 GACTTCAGAGTTCAGGAGCTGCCACGATATCGATTTCGCGTGGCGATACACTGTT 2643
Qy 522 -----SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIle--- 538
Db 2642 TACTATGGGAGTAGGGAGCGTTATATCCGGCA---TCGGCTCTCACGCGTGGTGTCC 2586
Qy 539 -----LeuProAspGluSerGluAsnGluValIleAspLeu--- 550
Db 2585 GTTACACCGGAGCTTTTTCGCGAGGATGATGTGAACGCTCTTATATCAAGCTGAAT 2526
Qy 551 -----ProGluArg 553
Db 2525 GCCTCTCCCTATCTTCGGAACGTGAATATCCGATTTCGTGGAGCACAAATGATAGGATGAG 2466
Qy 554 ThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeu 573
Db 2465 ATTGCTCTGGCGGAT-----AGCTCTCGCCCTGTGGACTGC 2430
Qy 574 TyrValPheValAlaSerAspLysProArgAspGlyCln-----IleGlyLeuGly 590
```

```

Db 2429 TATATTCTTACCGTTCGGCCCAAGAGCAAAATCGTTTCGAAGCGCGAAGTCTCTCGGCAACCAAT 2370
      |||...  . ||| ...  :|||
Qy 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPhe--GluHisAsnLeuLeuAsn 609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2369 TCCGCTGGAGACTTCGGGGCGCTTTGTCTCTCGGTTTACACGATCGCAATTTGTTTGGT 2310
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 610 -----ArgAspGlyTyrGlnAlaGlyValAlaGluLeuArg 620
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 2309 GGGCGGAGAGTTCATATCAAACTCAAGGTGCTTACGAAGCAATTCGCAAGGGTTCG 2250
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysPro----- 635
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2249 CACAGCTTCATGGAATATGGGTGGAAGCTCGCTCCGTTTCCCTCGTCTCCTCTCCCA 2190
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 636 ---LeuSerHisProLeuAsnAspGlnAlaThrArgAlaThrLysGlnGlnGluVal 654
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 2189 TTCATTCTGACGAACCGCGCGGCTACGGGATCCACGGAATGGAAG----- 2139
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 655 PheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluLeuSer 674
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2138 -----ATCGGTATATAATACCAGACACGTCGC-----GAGTTTGAT 2103
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 675 ArgSerIleIleGln-----AsnGlyGlyTyrAsnArgThrTyr-----Ser 688
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2102 CGGTGATCTCTCGCTCACTCAATTAATTCATGG---CAGACCTACTCGCACAAATCGT 2046
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 689 LeuArgTyrArgLeuAspLysThrGlnAlaProGluThrTyrGlnAspLeu 708
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2045 CTGGTCATACGATCCGCTGCTGATCGAT-----TATCTCCATCTC 2001
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 709 ProVal-----AspPheValAsnGlyLysProSerGlnGluAlaLeu----- 722
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2000 CCGTATACATCGATCCCGACTTCGCCCAATCCCTTCGCGCTACGACTGCATGTATAACTAC 1941
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 723 -----LeuAlaGlyValAla-----ValHisLysThrValAlaAspAsnLeu 736
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1940 ACGGAGCAGTTTATCTCGCTCGGCATATATCTGTAATATACCATCCGCTTCGTCCTATG 1881
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 737 -----ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGly 752
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1880 GAGCGTACCGTATCCAACTCT-----TTTACGCGCAGCGTTTCAGTATCCAGACAGCC 1830
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 753 SerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyVal 772
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1829 GGCAACCTGCTGCAAGCCATTTCTTATCTGACCGGATTTCCGAAAGACGACAGCGGTG 1770
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 773 TyrSer---PheGly----- 776
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1769 TATAAAATGTTGGTCTGCACATATCTCAGTTCGTCAGCTCGATCTCGATCGGCTAAA 1710
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 777 -----AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly 790
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1709 ACCGTTCTTCGAAAGAGCAATACTTTG-----GCACCTGCATCTGGGTTC 1662
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 791 GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArg 810
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1661 GGACTGGCTTTCCCTTAT-----GGCAATGCTCGCCATATACCCCTTTGAGTTAGCT 1611
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 811 PhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIle 830
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1610 TACTTTGGCGGAGGATCGAACAGCGTTTCGGGCTGGAGTGTCCGATCCCTCGGCCCGGGG 1551
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 831 Ser-----AspLysGlyTyrLeuThrGly---GlyGlnValLeuAlaVal 844
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1550 ACTATGAAGATGACTCCCGACAGACCTTCTTCGATCAGATGGGTGATATTCCTCTGGAT 1491
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 845 GlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAsp 864
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 CTGAATGTCGAATACAGGACAAAGCTGTTCTGGAAGTTTCGGCGCAGCAGCTTTTCTCGAT 1431
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 865 IleGlyAsnAla-----Tyr 869
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1430 GCCGCAATGTCTGCAGATAAAGGAGTATGAGAATCAGGAGCGGTCTCTTTGGTTTC 1371
Qy 870 AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1370 GATCCG---TTCACAGAATAAGTATGGCTACGGTCTGGGCTTCGTCGTCGACTTC 1314
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLys----- 903
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1313 GATTATTTCTGTGCGCGTGGAT-----GCCGACTGAAAGCCTACGATCCTCAGCAG 1260
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 904 -----GluGluGlyAsnProIleLysLeuHis 912
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1259 ACAGGCGCTTCAAAATGGCTATACACGCCCAACCTTTCTTCCAATTTCTGCTTGGCAC 1200
Qy 913 PhePheIleGlyThrProPhe 919
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1199 ATTGCAGTAGGCTATCCGTTTC 1179
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
RESULT 34
US-08-961-083-117
; Sequence 117, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-117
Alignment Scores:
Pred. No.: 0.00035 Length: 3121
Score: 134.00 Matches: 190
Percent Similarity: 32.04% Conservative: 124
Best Local Similarity: 19.39% Mismatches: 318
Query Match: 2.83% Indels: 348
DB: 3 Gaps: 53
US-09-914-168-2 (1-919) x US-08-961-083-117 (1-3121)
Qy 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIle-----IleAsnHis 41
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 155 TTGACGACCGCAGATTTTATCTGCTCCTATAATAGTACGCTTTCTATCGGAGTCGGGAACAT 214
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

Qy	42	ValPro-----AlaHisAspThrAlaIleAsnGlnAlaLys	53
		:::	:::
Db	215	FTACCAGAGCCTCTGAANAATCGAAGGTTATCAATATATGGTTATATATGAACTAAAGAA	274
Qy	54	AlaGlyAsnProProValLeuLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaLa	73
Db	275	CAGGATAATACAGAGCTTCAAGGACA-----GTTGATGGAAATACTCTGCTCA	325
Qy	74	GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAsp	91
		:::	:::
Db	326	AGAGTAGTCAACCAACTCTCAAAACATGATGATGTTATTCACCTGATTAGAA	385
Qy	92	-----GlnSerProIleSerArgIleGlyGluGlnSerProLeuGlyLeu	107
Db	386	TGGAACCAAGACAGCGGAAGTTAGTTTCAAGGTTGAAGCATCAGGGCATGGACTT	445
Qy	108	AspMetSerValIleGluGluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGlu	127
Db	446	-----TCAGAAAATCTCTPATAGCAGCAGACAATCTATCTTCTAATGAT	490
Qy	128	SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro	147
Db	491	TCATTCCGAAGTCAAGTTGAGCAGAA-----CCGGATCACAAAGGAAATCT---	538
Qy	148	AsnSerGluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArg	167
Db	539	-----GTAGTTCGACCAACAGTG---CCAGAACAAGGAATCTCTGTCTGCT	583
Qy	168	LeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPhe	187
Db	583	-----	-----
Qy	188	TyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGlu	207
		:::	:::
Db	584	---ACAACGGTGCAGATGCGGAAGAGGAAGTATTGGCAGCAGCAAAATGATCGACAG	640
Qy	208	ProTyrAlaAsnIleLysAlaAlaLeuGluAspIle---ThrGlnLysSerAlaMetAsp	226
Db	641	-----TATAACTCTCCATTGGAAACCAAGGCAGCCAGAACCCGGTCATGG	688
Qy	227	LeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGly	246
		-----	-----
Db	688	-----	-----
Qy	247	TyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIle	266
Db	689	-----GGTGAAGCC---GCAGTCCGT	706
Qy	267	HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly---Glu	285
Db	707	GAAGACTTA-----CCAGTCTACACT-----AAGCCACTAGAAACCAAGGTACACAA	754
Qy	286	GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAsp	305
Db	755	GGACCCGGCATGAGGTGAGCTGCGAGTTCCGAGGAAGAACAGCTTACACAGAACCG	814
Qy	306	ValPhe-----HisHisGlyLysTyrGluThrLysLysAsn	317
Db	815	TTAGCAACGAAAGCAGCAAGAGCCAGGTCATGAGGCAAAGCTACAGTCCGCGAAGAG	874
Qy	318	LeuIleGlu-----AsnAlaSerAlaGluHisGlyTyr	328
		:::	:::
Db	875	ACTCTAGGTACAGGAACCGGTAGCGCAAAAGGCACACAAAGCCGGAACAT-----	928
Qy	329	PheAspGlyArgTrpLeuAspSerValAspValIleLeuProAspAsnThrAlaAsp	348
Db	929	-----GAGGCG-----GAACGGSCAGTAGAAGAAGACTTCG-----	961
Qy	349	ValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThr	368
		:::	:::
Db	962	---GCTTTAGGTGCTACTACACGAAATAGACGGAATCCAGAATATT-----	1006

[illegible]

Db 1664 ATTACA---GAAATACGATGGAAACGTATATAAAGTAGCGGTGATCAACTTGTCT 1720  
Qy 733 AlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGly 752  
Db 1721 GAAGAAGGT-----ACAGCGGTTCACAGATGATTTACATCTTCTGTGACT 1768  
Qy 753 SerSer-----GlyLeuValSerAspAlaAsnMetAlaIleAla----- 765  
Db 1769 AAATCTAAAGCAGACGACGAGGAGTTTACACATCCTTTAAACAGCTGGTAACGCCATG 1828  
Qy 766 ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsn-----Ala 779  
Db 1829 CAAAGCAATCTGTCTGTCTATACATCTGCTTACAGATATGACCGCAGATGAGTCAAGC 1888  
Qy 780 TyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSer 799  
Db 1889 TTAGCGCATTAAGCAGACGAGTATCTCACAGGTCCATTACAGCGAGCTTCATCGGTCT 1948  
Qy 800 AspAsnPhe---AsnHisValProTyrArgLeuArg-----PhePheAla 813  
Db 1949 GATGGAACAAATCGTATGCCATTATGATTTGAAGAAACCATTTATGATACATTAAT 2008  
Qy 814 GlyGly-----AspGlnSerIleArgGlyTyrAlaHisAspSer----- 826  
Db 2009 GGTCTACAGTTAGAGATTTGGATATTAAACACTGTTCTGCTGATAGTAAGAAAAATGTC 2068  
Qy 827 -----LeuSerProIleSerAspLysGlyTyr 835  
Db 2069 GCACCGCTGGGAGGAGCGAATAGCGCGAATATTAAATATGTTGCGAGTAGAAGGAAA 2128  
Qy 836 LeuThrGlyGlyGlnValLeu-----AlaValGlyThrAlaGluTyr 849  
Db 2129 ATCTCAGGTGCGAAATCTGTTGGCGGATTAGTAGCGAGCGCAACAAATACAGTATAGAA 2188  
Qy 850 AsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr 869  
Db 2189 AACAGCTCGTTTACAGGGAAACTT-----ATCGCAAAATCACCAG 2227  
Qy 870 AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGly-----ValArgTyr 887  
Db 2228 GACAGTAAT---AAAAATGATACT-----GGAGGAATAGTAGTAATATACAGGAAT 2278  
Qy 888 AlaSerProValGlyGlnValArgValAspValAlaThrGlyValLysGluGlyAsn 907  
Db 2279 AGTTTCGAGAGTTAATAAAGTTAGGTAGATGCCTTAATCTCTACTAATGACACGCAATAAT 2338

RESULT 35

US-08-770-379-20  
; Sequence 20, Application US/08770379  
; Patent No. 5849564  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A.  
; APPLICANT: Russo, James J.  
; APPLICANT: Edelman, Isidore S.  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/770,379  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Alignment Scores:  
Pred. No.: 0.0302 Length: 32207  
Score: 132.50 Matches: 167  
Percent Similarity: 33.29% Conservative: 118  
Best Local Similarity: 19.51% Mismatches: 300  
Query Match: 2.80% Indels: 271  
DB: 2 Gaps: 39

US-09-914-168-2 (1-919) x US-08-770-379-20 (1-32207)

Qy 6 LeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeuProLeuMetThr 25  
Db 1514 CTGTATCTCCGAACTATTATCAGCGTGACCTTA----- 1549  
Qy 26 SerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIleLeuAsnHisValProAlaHis 45  
Db 1550 -----CACAGTACAAACCTTGGAAACCTTAGTTACCAAGCTACCA----- 1588  
Qy 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGln 65  
Db 1589 -----AAACTTGTGTAGGCGTTCACCGGGCGCCGACCTCTCTGACAAAGCCG----- 1636  
Qy 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85  
Db 1637 ---TCCGTGAGACAGTCCCTTCCACCTGTGTCACAAACCTGTCGCGAGATGCGCTGGAC 1693  
Qy 86 ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeu 105  
Db 1694 GCCCTG-----GAAAAAAGGATCCG--- 1714  
Qy 106 GlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGluPheAla 125  
Db 1715 -----GCCCTTCTTGTGAGGGGACCACGCTGGCGCTGGAGACACTCTCTAGGA 1762  
Qy 126 GlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGlu 145  
Db 1763 TACGGGTCTGTGCAG-----GACTACAAAGGAGACGGTACAG----- 1798  
Qy 146 GlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeuIle 165  
Db 1799 -----ATAATATCCAGCCTTGTGGGCATCCAAAA-----TTAGTCT 1834  
Qy 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla 185  
Db 1835 AGGACACCGGCGCGACAGTGGCCACTGCCGTGACAAAGCTAACTGACCTCAATCA 1894  
Qy 186 -----LysPheTyr----- 188  
Db 1895 ACTCTGCCACGACCGCCATCAGACGGCTACGAAACGAAATACATAGATTGATCA 1954  
Qy 189 -----GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys 205  
Db 1955 AGGACCTCAAGAGGCTCAAAAACACGAGACC-----AATCGGCC 1996

Qy	206	ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMet	225
Db	1997	ATGGAGGAATGGAGCAGAGAAGTACTTGGCTCTTGACATGGCTCTCCGGAACGTGTCCGC	2056
Qy	226	AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArg	243
Db	2057	ACCTCTC-----CTGCACACAGGCTCCACCGCGAAGCTTAGAGATTT	2098
Qy	244	AlaValGlyTyrTyrAspIleAspLeuSerIle-----IleArgAsnSer	258
Db	2099	GCAGAGAAGCACITTCANAATACTACTCCCGCTACCCGGGACGCCCGCTCAACAGCTCT	2158
Qy	259	IleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg	278
Db	2159	CCAAAGCGCGATGGAATACAGCGCACGCCCTCCCGGACCCAAAGGATATAGAC---AGA	2215
Qy	279	AlaValGluValArgGlyGluGlyAlaAspAspLys-----	290
Db	2216	GCTACATCCATCCACGGGGAACAGGCGTGGAAAGAAGATACACAGAGCGCTTCAAGGATTTC	2275
Qy	291	-----AlaPheThrThrValAlaAsp-----	297
Db	2276	AACITTCGCGCTTCGGCGCCGCTGACTGGATGCCCTGGCAGCGGAGTACCACGCCGT	2335
Qy	298	-----GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLys	315
Db	2336	GGTTACGCCCTTCCGGCGCGCTGGGTCCACGGCGCTC---TCAGGGTTCTCTGGAGACGATC	2392
Qy	316	LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp	335
Db	2393	CTAGGGAGCGCTGAACGACATC-----TACATGGAT-----	2422
Qy	336	ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr	355
Db	2423	--AAGCTCCGCTCCTTCTGCCCCGAGCGCAGCCCTTTTCAGCGCCGCCCTTCGACTGG	2479
Qy	356	GlyThrGlnTyrArgPheAspGluValValPhePhe-----ThrIleAspProLysThr	373
Db	2480	CTAAGCGCGTATCAG---GACCAGGTACGCTTTTCTTTCGCACCATCAGG-----	2527
Qy	374	AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeu	393
Db	2528	-----CTGCGCTGGTGGCAGCGCTGGCGGACCAAGATC	2560
Qy	394	LeuThrValAsnMetGlyGluAlaTyrAsnLeu-----GlnAlaVal	407
Db	2561	AGCGTGCAGGCACTGAGGCTTACCCAGCGCTCCAGTCCGGCGATTTCACGAGCGCCACG	2620
Qy	408	ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetVal---AsnThrGlu	426
Db	2621	GTGGGCACGCCCTTGGAGCTTCCCTGTCACACAGATAGCGCGCATCGCTCCCAACATGAAG	2680
Qy	427	IleValPheProGluArgGlu---GlnIleGlnAsnAspGlnValSerPheGluGlnSer	445
Db	2681	TCCGTGTTTCAACGACCACGAGCTTACAGTTCGATCGCATAGGTCGGGATTATGTGGAGGCC	2740
Qy	446	SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu	465
Db	2741	CAACGACCGGACGACACGCCACACGCTCCCGCTCCAAAGATACAGGACCAAAAGACT	2800
Qy	466	ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer	485
Db	2801	CTGATTCACATCCGGACGAATCGTCGGGACGAGTACCC---GCCTTTCTTAAAGACG	2857
Qy	486	AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet	505
Db	2858	TCCCTACTGCAGCAAGGACCAACTTCTGGCGCTACAG-----	2896
Qy	506	ProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly	525
Db	2896	-----	2896
Qy	526	ArgIleSerAspAlaValSerAlaValAlaAlaArgAlaIleLeuProAspGluSerGluAsn	545

D	b		2897	- - - - -	- - - - -CGGCGCAGCTTC	2908
Q	y	546	GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal	565		
D	b	2909	GAGTGCgtcagcgacgacatgcgccggcgagaggccagagaagcatcgccgagaa	2968		
Q	y	566	TyrGlnSerLys-	-- -- --LysValproLeuTyrVal	575	
D	b	2969	ACCCAGCGCAAAATGGCACACGCCCATCTCAGCTCTTTACAGCAGGCACCAGTGCGCATC	3028		
Q	y	576	PheValAlaSeraspLysProArgaspGlyGlnIleGlyLeuGlyTrpGlySerAspThr	595		
D	b	3029	- - - - -TCGGGGCCCCGCTATCCTTACAGACCCCGGTGGGCTTC-	-- -- --	3067	
Q	y	596	GlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrglnAla	615		
D	b	3068	- - - - -CTCAGAGGCATCATACGACAAGTCTCTGGAGCGGAATTCTACGAGACG	3118		
Q	y	616	GlyAlaGlu- - - - -LeuArgLeuSerGluAaspLysLysGlyValLysLeuTyrAla-	-- -- --	632	
D	b	3119	GGTCTCAGAGGACTGTCTCGCTCGACAGACCATCAAGCTCCATCACCCTATACGCTCCC	3178		
Q	y	633	- - - - -ThrLysProLeuSerHisProLeuAsnAspGlnLeu-	-- -- --ArgAlaThr	647	
D	b	3179	GTAGAGAGAAGCAAAGTAATGCAGCTGCTGCTGCAGCAGGTGAAAAGACGACGACAAAC	3238		
Q	y	648	LeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg	667		
D	b	3239	ACTGAGACCGCTCTCGAGCTAGAGCGCGGCTTACGCACGCGCAGAC-	- - - - -GCTAGA	3292	
Q	y	668	ThrLeuGluHisGluIleSerArg-	- - - - -SerlleGlnAsnGlyGlyTrpAsnArg	685	
D	b	3293	CTCCTGACGAGCGGTGCGATGAGCTGTCCACGCTTCGCGGTTAAGGGGGGAAGCCCGCG	3352		
Q	y	686	ThrtyrSerLeuArgTyrArgLeuAaspLysLeuLysThrGlnAlaProGluThrTrp	705		
D	b	3353	GTGGAATCCTGGCGGCAGAAAAATCCAACCTCGMAATCCCTCGTA-	- - - - -	3397	
Q	y	706	GlnAspLeuProValAspPheValasnGlyLysProSerGlnGluAlaLeuLeuAlaGly	725		
D	b	3398	- - - - -CAGGAAGCGGACGAGCGCCGC	3418		
Q	y	726	- - - - -ValAlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyTyr	743		
D	b	3419	CTCCTGTGTGGCCACCATAGACAGGTGGCC-	- - - - -	3488	
Q	y	744	ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla	763		
D	b	3449	- - - - -GGCAGGCCCCAGGACGACC	3466		
Q	y	764	IleAlaArgAlaGlyllleSerGlylvalTyrSerPheGlyAspAlaAla	779		
D	b	3467	ATATCACCATTCCACACTCCAGGGACTIONTACCAACAGGGACGAGAGGCC	3514		
RESULT 36						
US-08-757-669A-20						
; Sequence 20, Application US/08757669A						
; Patent No. 6183751						
; GENERAL INFORMATION:						
; APPLICANT: Chang, Yuan						
; APPLICANT: Bohenzky, Roy A.						
; APPLICANT: Russo, James J.						
; APPLICANT: Edelman, Isidore S.						
; APPLICANT: Moore, Patrick S.						
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS						
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF						
; NUMBER OF SEQUENCES: 20						
; CORRESPONDENCE ADDRESS:						
; ADDRESSEE: Cooper & Dunham LLP						
; STREET: 1185 Avenue of the Americas						
; CITY: New York						
; STATE: New York						

RESULTS 36

```

US-08-737-669A-20
: Sequence 20, Application US/08757669A
: Patent No. 6183751
: GENERAL INFORMATION:
: APPLICANT: Chang, Yuan
: APPLICANT: Bohenzky, Roy A.
: APPLICANT: Russo, James J.
: APPLICANT: Edelman, Isidore S.
: APPLICANT: Moore, Patrick S.
: TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
: TITLE OF INVENTION: SEQUENCES AND USES THEREOF
: NUMBER OF INVENTIONS: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York

```



COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/757,669A  
 FILING DATE:  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 45185-F  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 278-0400  
 TELEFAX: (212) 391-0525  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 32207 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-757-669A-20

Alignment Scores:  
 Pred. No.: 0.0302 Length: 32207  
 Score: 132.50 Matches: 167  
 Percent Similarity: 33.29% Conservative: 118  
 Best Local Similarity: 19.51% Mismatches: 300  
 Query Match: 2.80% Indels: 271  
 DB: 4 Gaps: 39

US-09-914-168-2 (1-919) x US-08-757-669A-20 (1-32207)

Qy	186	-----LysPheTyr-----	188
Db	1895	ACTCTGCCACGACCGCATCGACGCGCTACGAAACGAACTATACAGATTGATCAAA	1954
Qy	189	-----GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys	205
Db	1955	AGGACCTCAAAGAGGCTCAAAACACGAGACC-----AATCGGCCC	1996
Qy	206	ThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGluSerAlaMet	225
Db	1997	ATGGAGGAATGAACCAAGTACTGCTCTTCAATGCTCTCCGGAACGTGTGCC	2056
Qy	226	AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAarg-----	243
Db	2057	ACCCTC-----CTCAACAGGCTCCACCGCGAGGCTTAGAGATT	2098
Qy	244	AlaValGlyTyrTyrAspIleAspLeuSerIle-----IleArgAsnSer	258
Db	2099	GCAGAGACGACTTCAAATACTACTCCCGTACCCGCGGACGCCCGCTCCAAGCGTCT	2158
Qy	259	IleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg	278
Db	2159	CCAACGCGGATGAATACAGCGCCAGCCCTCCCGGACCCAAAGGATATAGAC---AGA	2215
Qy	279	AlaValGluValArgGlyGluGlyAlaAspLys-----	290
Db	2216	GCTATCTCCACGCGGAACGCGTGGAAGAAGATACAGCAGCGCTTCAGGATTTC	2275
Qy	291	-----AlaPheThrThrValAlaAsp-----	297
Db	2276	AACCTTCGCGCTCTGCGCGCGCTGACTGGGATCCCTGGCAGCGGAGTACCAACGCGT	2335
Qy	298	-----GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGlnLys	315
Db	2336	GGTTGCGCCCTTCGCGCGCGTGGTCCAGCGCTC---TCAGGTTCTCTGAGACGATC	2392
Qy	316	LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp	335
Db	2393	CTAGGCGAGCTGAACGACATC-----TACATGGAT-----	2422
Qy	336	ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrFasThr	355
Db	2423	---NAGTCTCCGCTCTTCTGCGCGCGCGAGCTTTTCAGGCGCCCTTCGATGG	2479
Qy	356	GlyThrGlnTyrArgPheAspGluValValPhePhe-----ThrIleAspProLysThr	373
Db	2480	CTAACGCGTATCAG---GACCAAGTCAGCTTTTCTTTCGCGCACCATAGG---	2527
Qy	374	AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeu	393
Db	2528	-----CTGCGCTGCTGCGAGCGCTGGCGGATATATGTGAGGCC	2560
Qy	394	LeuThrValAsnMetGlyGluAlaTyrAsnLeu-----GlnAlaVal	407
Db	2561	AGCTGCGAGGCACTGAGGCTTAGCCGCGCTCCAGTCCGCGGATTTGACGAGCGCCACG	2620
Qy	408	ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetVal---AsnThrGlu	426
Db	2621	GTGGCGACGCCCTTGAGCTCCCTGCCACAGAGTACGCGCGCATCGCTCCACATAGAG	2680
Qy	427	IleValPheProGluArgGlu---GlnIleGlnAsnAspGlnValSerPheGluGlnSer	445
Db	2681	TCCGTGTTCACAGCACCAGGACTTTCAGTGGGATCAGAGGTCGCGGATATATGTGAGGCC	2740
Qy	446	SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu	465
Db	2741	CAACGACCGGACGACACCGCTCCACGCTCCAAAGATACAGGACCAAGACT	2800
Qy	466	ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer	485
Db	2801	CTGATTCACATCCGCGGCAATCTGCGCGGACGAGTACCC---GCCTTTCTTAAGACG	2857
Qy	486	AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet	505

```
Db 2858 TCCCTACTGCACAGAGCCCAAACTTGGCGCTACAG----- 2896
QY 506 ProAspArgValLeuAlaIleAsnHisAspGlyValAsnArgSerIleLeuGly 525
Db 2896 ----- 2896
QY 526 ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn 545
Db 2897 -----CGGCGCGGACTTC 2908
QY 546 GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal 565
Db 2909 GAGTCGCTCGAGGCGACATCGGCCCGCAGAGGCCAGAGAAAGCATCCGCGGAGGAA 2968
QY 566 TyrGlnSerLys-----LysValProLysVal 575
Db 2969 ACCCAGCGCAAAATGGACACGCCATCACTCAGCTCTTACAGCAGCGCACCCAGTCGCATC 3028
QY 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThr 595
Db 3029 -----TCGGGGCGCCGCTATCCTTACAGGACCGCGTGGCTTC----- 3067
QY 596 GlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla 615
Db 3068 -----CTCAGGAGCATCATATACGACAAAGTCTCTGGAGCGCGAATCTCAGGAGCG 3118
QY 616 GlyAlaGlu-----LeuArgLeuSerGluAspLysGlyValLysLeuTyrAla--- 632
Db 3119 GGTCTCAGGAGGACGTCTCGTCTGAGCAGACCATCAAGTCCATCACCCTATACGCTCCC 3178
QY 633 -----ThrLysProLeuSerHisProLeuAsnAspGlnLeu-----ArgAlaThr 647
Db 3179 GTAGAGGAGAGCAAAAGTGCAGCTCTCTGTGACGAGGTGAAAGACGACGAGCAAC 3238
QY 648 LeuGlyTyrGlnGlnIleValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg 667
Db 3239 ACTGAGACCGCTCTCGAGCTAGAGCGCGCTACGACGCGCAGCAGAC-----GCTAGA 3292
QY 668 ThrLeuGluHisGluIleSerArg-----SerIleIleGlnAsnGlyGlyTyrAsnArg 695
Db 3293 CTCTCGCGGACGGTGCATGAGCTGTACCGTTCACCGTTCGCGTTAAGGGGGGAAAGCGCGC 3352
QY 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrp 705
Db 3353 GTGGAATCCTGGCGGAGAAAATCCAAACCTCAATCCCTGTGA----- 3397
QY 706 GluAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGly 725
Db 3398 -----ValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
QY 726 -----ValAlaValHisLysThrValAlaAspAsnLeuValSerAspAlaAsnMetAla 763
Db 3419 CTCCTGTGGCCACCATAGACAGTGGCC----- 3448
QY 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 3449 -----GGCCAGGCCCCAGGAGACC 3466
QY 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAla 779
Db 3467 ATATACCATCCACATCCAGGACGTGTACCAACAGGAGGAGGCC 3514
```

## RESULT 37

```
US-09-230-371A-20
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
```

```
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE OF INVENTION: US97/0371A
; CURRENT APPLICATION NUMBER: US/09/230, 371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20
```

## Alignment Scores:

```
Pred. No.: 0.0302 Length: 32207
Score: 132.50 Matches: 167
Percent Similarity: 33.29% Conservative: 118
Best Local Similarity: 19.51% Mismatches: 300
Query Match: 2.80% Indels: 271
DB: 4 Gaps: 39
```

US-09-914-168-2 (1-919) x US-09-230-371A-20 (1-32207)

```
QY 6 LeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeuProLeuMetThr 25
Db 1514 CTGGTATCTCCGAATCATTTATCACCGTGACCCCTA----- 1549
QY 26 SerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleAsnHisValProAlaHis 45
Db 1550 -----CACAGTACAAACCTTGAAACCTATTACCAACCTACCA----- 1588
QY 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrProGluGln 65
Db 1589 -----AAACTTGGTGAGCGGTTACCGGGGGCCGACCTCTCTACAAAGCCG----- 1636
QY 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
Db 1637 ---TCGCTGAGACAGTCCCTTTCCACCTGTGCACACCCCTCGCGAGATCCCTGGAC 1693
QY 86 ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProLeu 105
Db 1694 GCCTTG-----GAAAAAAGGATCCG----- 1714
QY 106 GlyLeuAspMetSerValIleGluThrThrProLeuSerLeuGluLeuPheAla 125
Db 1715 -----GCCCTTCTGTCAGGGGACCGTGGCGCTGGAGACACTCTTAGA 1762
QY 126 GlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlu 145
Db 1763 TACGGGTCCGTGCAG-----GACTACAAGGAGACGGTACAG----- 1798
QY 146 GlnProAsnSerGluValValProThrLeuGluProGluLysProGlyLeuIle 165
Db 1799 -----ATAATATCCAGCCTTGGCGCATCCAAAA-----TTAGTC 1834
QY 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla 185
Db 1835 AGGACACCGCGCGGACCAAGTGGGCCACTGCGGTGACAGGGCTAACTGACCTCAATCA 1894
QY 186 -----LysPheTyr----- 188
Db 1895 ACTCTGGCCACGACCGCATCGACAGCGCTACGAAACGGAATATACAGATTGATCCAA 1954
QY 189 -----GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys 205
Db 1955 AGGACACCTCAAGAGGCTCAAAACACGAGACC-----AATCGGCGCC 1996
QY 206 ThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGluSerAlaMet 225
Db 1997 ATGGAGGAATGGAAGCAGCAAGAGTACTGCTCTTCAATGCGTCTCCGGAACGTTGCC 2056
```

Qy 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArg----- 243  
Db 2057 ACCCTC-----CTGCAACAGGCTCCACCAGGCTTAGAGAGTTT 2098  
Qy 244 AlaValGlyTyrTyrAspIleAspLeuSerIle-----IleArgAsnSer 258  
Db 2099 GCAGAGAAGCACTTCAAATACTACTCCCTCCCGGGAGCGCCCGCTCCAAAGCGTCT 2158  
Qy 259 IleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg 278  
Db 2159 CCAACGGCGATGGAATACAGCGCCAGCCCTCCCGGACCCAAAGGATATAGC---AGA 2215  
Qy 279 AlaValGluValArgGlyGluGlyAlaAspAspLys----- 290  
Db 2216 GCTACATCCATCCAGGGGAACGGCGTGGAAGAAGATACAGCAGCGCTTCAAGGATTTC 2275  
Qy 291 -----AlaPheThrValAlaAsp----- 297  
Db 2276 AACTTCGCGCGTCTGGCGCCCGCTGACTGGGATGCCCTGGCAGCGGATACCAACGCCGT 2335  
Qy 298 -----GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLys 315  
Db 2336 GGTTCGCCCTTCGGCGCGCGTGGTCCAGCGCTC---TCAGGGTTCCTGGAGACGATC 2392  
Qy 316 LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp 335  
Db 2393 CTAGGGAGCGTGAACGACATC-----TACATGGAT----- 2422  
Qy 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355  
Db 2423 ---AAGCTCCCGCTCTTCTGCCCGACCGCGAGCGCTTTTCAGCGCCCGCTTCGACTGG 2479  
Qy 356 GlyThrGlnTyrArgPheAspGluValValPhePhe-----ThrIleAspProLysThr 373  
Db 2480 CTAACGCGGTATCAG---GACCAAGTCAGCTTTTCTTGGCCACCATAGG----- 2527  
Qy 374 AsnGlnLeuThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeu 393  
Db 2528 -----CTGCCGCTGGTGGCAGCGCTGGCGCGCAAGAATC 2560  
Qy 394 LeuThrValAsnMetGlyGluAlaTyrAsnLeu-----GlnAlaVal 407  
Db 2561 AGCGTGACAGGCACTGAGGCTTAGCCACGCGCTCCAGTCCGCGGATTTCCACGAGGCCACG 2620  
Qy 408 ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetVal---AsnThrGlu 426  
Db 2621 GTGGCAGCGCCCTGGAGCTCCCTGCCACAGATACGGCGCATCGCTCCCAACATGAAG 2680  
Qy 427 IleValPheProGluArgGlu---GlnIleGlnAsnAspGlnValSerPheGluGlnSer 445  
Db 2681 TCCGTGTTCACAGCACCGGACTTCAGTGGCATCAGAGGTTCGGGGATTATGTGGAGGCC 2740  
Qy 446 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465  
Db 2741 CAACGAGCGGACGACACACCCACAGCTCCACGTCCTCAAAGATACAGCCACCAAGACT 2800  
Qy 466 ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer 485  
Db 2801 CTGATTTCACATCCGGACGCAATCGTCGGCAGCGACTACCC---GCCTTTCTTAAGACG 2857  
Qy 486 AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet 505  
Db 2858 TCCCTACTGCAGAGGCGCAACTTCTGGCGGTACAG----- 2896  
Qy 506 ProAspArgValLeuAlaIleAsnHisAspGlyValAsnArgSerIleLeuGly 525  
Db 2896 ----- 2896  
Qy 526 ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn 545  
Db 2897 -----CGGCGCGACTTC 2908  
Qy 546 GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal 565

Db 2909 GAGTCGCTCGAGAGCGACATCGCGCGCCGACAGAGGCCAGAGAAAGCATCGCGCAGGAA 2968  
Qy 566 TyrGlnSerLys-----LysValProLeuTyrVal 575  
Db 2969 ACCCAGCGCAAAATGGCACACGCCATCACTCAGCTCTTTACAGCAGCGCAGTCGCGATC 3028  
Qy 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPheLysArgThr 595  
Db 3029 -----TCGGGCGCGCCGCTATCTTACAGGACCCGGTGGGCTTC----- 3067  
Qy 596 GlyThrArgLeuValThrLysPheGluHisAsnLeuLeuAsnArgAspGlyTyrGlnAla 615  
Db 3068 -----CTCAGGGCATCATATACGACAAAGTCTCTGGAGCGCAATCTCTAGCAGACG 3118  
Qy 616 GlyAlaGlu-----LeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla--- 632  
Db 3119 GGTCTCAGGAGACTGTCTCTGGCTCAGCAGACCATCAATCCATCACCCTATACCTCCC 3178  
Qy 633 -----ThrLysProLeuSerHisProLeuAsnAspGlnLeu-----ArgAlaThr 647  
Db 3179 GTAGAGGAGAAACAAAGAAATGCACGTCTGCTGGCAGAGGTGAAAAGCAGCAGCAAC 3238  
Qy 648 LeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg 667  
Db 3239 ACTGAGACCGCTCTCGAGCTAGAGCGCGCTACGCGCAGCGCAGC-----GCTAGA 3292  
Qy 668 ThrLeuGluHisGluIleSerArg-----SerIleLeuGlnAsnGlyTyrTrpAsnArg 685  
Db 3293 CTCCTGCAGCAGCGGTGCGATGAGCTGTACCGTTGCCGTTAAGGGGGGAAGCCCGG 3352  
Qy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrp 705  
Db 3353 GTGGAATCCTGGCGCAGAAAATCCAAACCTGAAATCCCTGGTA----- 3397  
Qy 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuGly 725  
Db 3398 -----CAGGAAGCGGAGCAGCGCGC 3418  
Qy 726 -----ValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743  
Db 3419 CTCCTGTGGCCACCATAGACACGGTGCC----- 3448  
Qy 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763  
Db 3449 -----GGCAGGCCCGCAGGAGCC 3466  
Qy 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAla 779  
Db 3467 ATATCACCATCCACTCCAGGACTGTACCAACAGGACAGGAGGCC 3514

## RESULT 38

US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: TUBERCULOSIS  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

## Alignment Scores:

```

Pred. No.:      211          Length:      4411529
Score:          131.50      Matches:       165
Percent Similarity: 33.21%  Conservative: 102
Best Local Similarity: 20.52% Mismatches:   276
Query Match:      2.78%    Indels:        262
DB:              4        Gaps:          36

US-09-914-168-2 (1-919) x US-09-103-840A-1 (1-4411529)

QY 15 ValAlaLeuAlaAlaTyrLeuProLeuMetThrSerGlnAlaLeuAlaGlnAlaAsnAsn 34
    |||||||
DB 3496875 GTCCCTTGGCGGCGCCATACCGCGCAGCACCGATCGCGCGCTACAGTTGCTGG 3496934
QY 35 ProAlaAsnIleAlaAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAla 54
    |||||
DB 3496935 CCG-----GTCCAGAGGTGACATCCCGTTGTCGCGCGCAAGAGCA 3496976
QY 55 GlyAsnProProValLeuLeu----- 61
DB 3496977 GCCCGTGGCGGGGTCAATGCTGCGCGCAGGTGAGTGCCTAGACCGCAATGAAAGCCATTGTG 3497036
QY 62 -----ThrProGluGlnIleGlnAlaArg----- 69
DB 3497037 GCGCAGGCTGCTTGGACACAGTGGACCGAGGGA-ATATCTCGCGGAGCTACCAATTGG 3497095
QY 70 -----LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
    |||||
DB 3497096 GAGTGGGCGCTAGGTTCAGTGGCGGGTTCGCCGCCAAGCAGCGCCATCAGACCC 3497155
QY 87 ValAsnPheAspGlnSerProIleSerArgIle-GlyGluGlnSerProProLeuGlu 106
    |||||
DB 3497156 CAGTGGCCCGATCCCGGTCGCTGTCGCGGGCCGCTGCTCTCAGCCGCTCC----- 3497210
QY 106 yLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGluPheAlaGlu 126
    |||||
DB 3497211 ----GACGTCTTACCGCTGACGACGCGCGCGATC----- 3497243
QY 126 nGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlu 146
    |||||
DB 3497244 -----CTGGCAGTGGCAGCAGCAGACAGCAGCAGCGGTGGCGCGCGCGGGA 3497296
QY 146 nProAsnSerGluValValValPro-----ProThrLeu 157
    |||||
DB 3497297 GGCGCGCAGCATCGTCTTGTTCAGCGCAGCGCAATGGGGCTGGCGTGTGCCCATCAC 3497356
QY 157 uGluProGluLys----- 161
DB 3497357 CGAACCGCTGGAGATCGCCAAAGCCGCGACCGCGTCCGTGCGGAGTGTTCGCGCGCGG 3497416
QY 162 -----ProGlyLeuIleLysArgLeu---TyrAlaArgLeuPheAsnAspGlyValAs 178
    |||||
DB 3497417 CGGTACCCCGCAGATGCTGCTCGAGTGGTGGCAGCGATCAATGCGGACCGCTGCC 3497476
QY 178 nLysValProArgLeuLysAlaLysPheTyrGlnSer-SerGlnSerGlyGluThrSerA 198
    |||||
DB 3497477 ACCGACGCCACGG-----CGCGAACTGTCCAGGTCTGTGAGTGGCGCGGAGAGCT 3497527
QY 198 laIle-----GlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaL 216
    |||||
DB 3497528 ACTGGCAGACAGCGGTGTGACCATCGACGACTGTTCGCTCGCG----- 3497571
QY 216 euGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgG 236
    |||||
DB 3497572 -----CCCGGTACGCTCGGAGGTGAATT 3497596
QY 236 lnThrAlaLeuValAlaAlaArg-AlaValGlyTyrTyrAspIleAspSerIleIle 255
    |||||
DB 3497597 GCCCGCGCGCGCTGCTCTCGCGCTGCCGAGGT-----TCGTTAAGCGCGTTC 3497644
QY 256 Arg-AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIle 275
    |||||
```

```

DB 3497645 CGGTGAACCTCGTCGGCGCAGCCCGCGCGCTTGTTCGGTACCTCGATGCACAACTCGTCT 3497704
QY 275 eAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrVa 295
    |||||
DB 3497705 CGACTTT-----GACCCGGACGGTCAACGCTGCTGCCCTTCGCATCGCCGACCGCT 3497755
QY 295 lAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLy 315
    |||||
DB 3497756 TGTGTACCGCTTCCCG-----AACCCAC-----CG 3497779
QY 315 sLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgThrLeuAs 335
    |||||
DB 3497780 CCTCGCCTGATCGCGACGCGCTGCTCGACCAC-----CGACAATGGACCCA 3497827
QY 335 pArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspTh 355
    |||||
DB 3497828 CGAA-----TTCAACGC 3497839
QY 355 rGlyThrGluTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGl 375
    |||||
DB 3497840 TGTGCGCNAACCCGAGTGGCGCAATTTGGCTACGCGCGCATC----- 3497882
QY 375 nLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu-----GlnLeuLe 394
    |||||
DB 3497883 -----GATTCTGCTCCGAGCGAGTATACCTCGCATGCTCCGTCCA 3497926
QY 394 uThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLe 414
    |||||
DB 3497927 GGTCAATAATGGTGGTCCGGATTTC----- 3497951
QY 414 ulleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGl 434
    |||||
DB 3497952 -----CTGTATAACGCTCTG-----CAGATCTCTACCACTCCGAGAGTC 3497992
QY 434 nIleGlnAsnAspGlnValSerPheGlu-----GlnSerSerSerSerAr 449
    |||||
DB 3497993 GTTCTGTCACCTTCAGGATTAGCTTCGTGCGGGACAGCACCCGTGCAAGCCAGGCAATCG 3498052
QY 449 gThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLe 469
    |||||
DB 3498053 CGAAGAGCG-----CTGGATGACATGGTTCATGGAGGTTCACGGCGCATACGATCCCGGT 3498106
QY 469 uThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGl 489
    |||||
DB 3498107 CGGT-----CAGTACGTCTGCTGCGCATCCGACGCTGCGAAG 3498145
QY 489 nAspLysLeu-----AsnLeuValAlaAlaLysAlaArgHisLeuTyr----- 503
    |||||
DB 3498146 TGGCCAAATTGCCAAGCCAGCGCGCTTGTGCGGCAACGGCGCATCATCTCGAGTTGTT 3498205
QY 504 -----AspMetProAspAspArgValLeuAlaIleAsnHisAsp----- 516
    |||||
DB 3498206 CGTCGGTGAAGCGCCCTGGACCGCTTGACTCAGCACAACACGACACCCGCTACGGTAC 3498265
QY 517 -AspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaAr 536
    |||||
DB 3498266 CTCTGGCGCGCAGCGCAACAGCAGCGCGGACCTCGCTCGCGCAGTTCGTCAGGCGCTT 3498325
QY 536 gAlaIleLeuPro-----AspGluSerGluAsnGluVal---IleAspLeuProGluArgTh 554
    |||||
DB 3498326 CCAATCGACCGCTGACCGCTCGCGGAATGCGGTTCAGCAAGACCTCCCGCAGCCG 3498385
QY 554 rAlaLeuAlaAsnArgLysThrPro-----AlaAspValTyrGl 567
    |||||
DB 3498386 CGCCCGCCCGCAATGTTTCGCGCCCAACAATGGAAGCCACAGCGCTGCCGACTGTTTCAA 3498445
QY 567 nSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgaspGlyGlnIl 587
    |||||
DB 3498446 TCACCAGCAGCTCCCC----- 3498461
QY 587 eGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLe 607
    |||||
DB 3498462 -----CAGTCAGCGCGGACGATCTCT 3498484
```

QY	101	GlnSerProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeu	---Ser	119
DB	6340	AATCAA	:	111
QY	120	LeuGluGluLeuPheAlaGlnGluSerThrGluMet	---GlyIleAsnProAsnAspTyr	138
DB	6385	ATAGATGATACGTTAGCACACACGCTACAAATATTAATGGTGC	CAATACAAATCGGTTA	6444
QY	139	IleProGlu	---TyrGlnGlyGluGlnProAsnSerGluValValValPro	154
DB	6445	GTGGATGAGAAATTTAGAACATGTTAGCAAAAGTTTACAACTG	TTGTTGTTGTTCAACTCAA	6504
QY	155		---ProThrLeu	157
DB	6505	ACTAAACACAAAGCTAAAGCAGACATTTGCTCAAGCAATAGG	TCAACAAAGGTCGACAATA	6564
QY	158	Glu	---ProGluLysProGlyLeuIleLysArgLeuTyrAla	170
DB	6565	GACCAGATCAAAATGCTACACAGNAGAAACAAGAGCCCTT	GAGAGACTTAATCAA	6624
QY	171	ArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla	LysPheTyrGlnSer	190
DB	6625	GAA	---ACAAATGGAGCTCAAT	6672
QY	191	SerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys	ThrGluProTyrAla	210
DB	6673	AATGTTACAGACGAAAAAATAATATATTAGAACCAATAA	GAAATGTGAACCTATTGTA	6732
QY	211	AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMet	Asp	226
DB	6733	ATTGTAACCAAGGCTAATGAAATTAATAGNAAAAAGCTCG	GACACAAACGACTTTA	6792
QY	227	LeuAsn	---GlySerIlePro	232
DB	6793	ATAAATCAAAATCAAGATCGACACTAGAAGAAAAACAATA	GACTTGGCAAAATTAGAA	6852
QY	233	ArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr	TyrAspIleAspLeu	252
DB	6853	GAGTAAGAAGTAACGCGTTAAATCAAGTA	---TCACAGCGACACTCAATAATGATGTG	6909
QY	253	SerIleIleArgAsnSer	---IleGlyGluVal	266
DB	6910	AAATTCGCGAAATTAATGGAATTCGTAAATTTCTGAGGTCC	ATCCTCGAGACTAATTT	6969
QY	267		---HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGlu	281
DB	6970	AAACGTAATGCTAAACAAGAAATTTGAACAAGATGCGCAAG	ATGCAAAATGATGACTATCAAT	7029
QY	282	ValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla	AspGluValProLeu	301
DB	7030	GCAAAATAATTAATCAACTAATGAAGAAAAATACGCGCTATA	---GATACAGATTAATGTA	7086
QY	302	LeuIleGlyAspValPheHisGlyLysGlyLysThrLysLysAla	LeuIleGluAsn	321
DB	7087	GCTAAATTTGATGCTATTACCAATATTAATATGCTACAACT	ACACAAATTAGTTAATGAT	7146
QY	322	AlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArg	SerValAspValIle	341
DB	7147	GCTAAAAATAGTGGT	---AACACGAGTATTGCCAAATA	7182
QY	342	LeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly	ThrGlnTyrArgPhe	361
DB	7183	TTACCAAGTACAGCAGCTCAAACTAATGCAATTTAGCAGCT	CTAGCTAGCGAAGCTAAAAAT	7242
QY	362	AspGluValValPhePheThrIleAspProLysThrAsnGlnLeu	ThrThrAspProAsp	381
DB	7243	AAAAACGCTATA	---ATAGATCAAAACACCAAAATGCGACAGCAGAAAGAAAGAA	7293
QY	382	LysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal	AsnMetGlyGluAla	401
DB	7294	GAGCAAAATAATAAGTTTGATCGCTTTCAAGAGACGACAGT	CTAATATCTCTTAAAGCG	7353
QY	402	TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla	ThrArgTyrPheAsn	421

[illegible]

7030	GC	AAATAAATAATCACTAATAGAGAAAAATACGCCCTATA	---	-GATAGAGTTAATGTA	7080
302	QY	LeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuLeuGluAsn	321		
7087	Db	GCTAAATTCATCTATTAAACAATATTACTAATAGCTACAACACTACACAATTTAGTTAATGAT	7146		
322	QY	AlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValaspValIle	341		
7147	Db	GCTAAATAATAGTGGT-----AACACGAGATTATGCCAAATA	7182		
342	QY	LeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPhe	361		
7183	Db	TTACCAAGTCACAGCAGTCAAAACTAATATGCATTAGCAGCTCTAGCTACGGAAGCTAAAAAT	7242		
362	QY	AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAsp	381		
7243	Db	AAAAAGCGCTATA-----ATAGATCAAAACACCAANTGCGACGACAGAGAAGAAAAAGAA	7293		
382	QY	LysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAla	401		
7294	Db	GAAGCAATAATAAGTTGATCGCTTTCAAGAAAGACGAGATCGCTATATCTCTTAAAAAGCG	7353		
402	QY	TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsn	421		

D0	/147	GCTAAATAGTGGT-----AACACGAGTATTATGCCCAATAA / 168
QY	342	LeuProAspAsnThrAlaAspValSerLeuLeuIleTyraSpThrGlyThrGlnTyArgPhe 361
		: : : : : : : : : : :
Db	7183	TTCACAGTACAGCAGCTCAAACACTATGCATTAGCAGCTCTAGTAGCGAAGCTAAAAAT 7242
QY	362	AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAsp 381
		: : :           : : : : : : : : :
Db	7243	AAAACGCTATA-----ATAGATCAACACCACAATCCGACAGCAGAGAAGAAAAGAA 7293
QY	382	LysLeuProValLysArgLeuLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAla 401
		: : :     : : : : : : : : :
Db	7294	GAAGCAAAATAATAAGTTGATCGCTCTCAAGAGAGACCATCTATATCTCTCTAAAGACG 7353
QY	402	TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsn 421

[illegible][illegible]

```
Db 7354 CACACTACTGATGAAGTTAATAATATATAAATCAAA-----GCTGTTCAAAATATTATTAAC 7407
QY 422 MetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSer 441
Db 7408 GCTGTTCAAGTTGCAAGTTATC-----AAGAAACAAACCGTTAAACACCAATTAATCAAA 7461
QY 442 PheGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461
Db 7462 TTCATTGATATCAAAAGAAATTAATGAA-----AATACGCTGATGCAACACTAGAA 7515
QY 462 ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro----- 479
Db 7516 GAAAAAGCTGAAGTAATAAGATTGCTTCAAAATGTACTACTTCCACATCAGATGAAAT 7575
QY 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuValAlaAlaLysAla 499
Db 7576 GCTAATGTAGATCATAACACAGAGGTGATCAAGCTTTAGATAAGCTAGACCAAAATC 7635
QY 500 ArgHisLeuTyrAspMetProAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
Db 7636 GAGGAATTTGACACAAAGTTAGTAAGAA----- 7665
QY 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539
Db 7666 ---CGAGATGTTTAAATGCAATCCAAAGCAATTTAATTCACAAACTCAAGAAATA--- 7719
QY 540 ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
Db 7720 -----CAAGAGAACCAAGAGCTACGAATGAAGAAAACTGAAGCATTAATATAA 7770
QY 560 -----LysThrProAlaAspValTyrGlnSerLysLysVal 571
Db 7771 ATRAAACCAATTACTTAATCAGGCTAAAGTAAATATTGATCAAGCACAGTCAAAATAAA--- 7827
QY 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrp 591
Db 7828 -----GATGTAGATAGTGCAGAA----- 7845
QY 592 GlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAsp 611
Db 7846 -----ACACGTAGTATTCAAGATATAGAG---CAATTCACACCATCCCA 7887
QY 612 GlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyr 631
Db 7888 CAAACAAAGCAACCGCGCTGACAGATTAATGAA-----AAAGCTAACCAACAACAA 7941
QY 632 AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGln 651
Db 7942 AGTACT---ATTGCAATCATCCT-----AATTCACAAATGGAAGAAGA 7983
QY 652 GlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671
Db 7984 CAGCAA-----GCAAGTGCAAAACTACAAGAA 8010
QY 672 GluIleSerArgSerIleIleGlnAsnGlyClyTrpAsnArgThrTyrSerLeuArgTyr 691
Db 8011 GTTCTTAAAAAGCAATAGCT----- 8031
QY 692 ArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGlnAspLeuProValAsp 711
Db 8032 AAATAGATAAAGTCAAAACCAAT-----GATGATGTAGAAAAGACT 8073
QY 712 PheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThr 731
Db 8074 GTAGTAAACGGAATGCTGAAATTTGAAATATATTACTGCTACT-----ACA 8121
QY 732 ValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluVal 751
Db 8122 GTTAAAGATAAAGCTAAGCTGATGTAATGCTGAAAGAGAGCAAAACCTACAAT 8181
QY 752 GlySerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGly 771
```

```
Db 8182 AATAGTAATGATCAAGCAAGCACTGAAGAAAAAATTAGTTGCT----- 8223
QY 772 ValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly 791
Db 8223 ----- 8223
QY 792 IleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisVal 805
Db 8224 -----AGTGACAATTTAATCACCCT 8244

RESULT 40
US-09-147-236-1
; Sequence 1, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiro
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 6537-011-0PCT
; CURRENT APPLICATION NUMBER: US/09/147,236A
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/JP97/03633
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 16836
; TYPE: DNA
; ORGANISM: Acetobacter xylinum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (869)..(1891)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3101)..(5368)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5373)..(7778)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7784)..(11761)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11764)..(12231)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12448)..(14652)
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or
; OTHER INFORMATION: c
US-09-147-236-1

Alignment Scores:
Pred. No.: 0.0162 Length: 16836
Score: 130.50 Matches: 198
Percent Similarity: 33.82% Conservative: 146
Best Local Similarity: 19.47% Mismatches: 367
Query Match: 2.76% Indels: 307
DB: 4 Gaps: 51

US-09-914-168-2 (1-919) x US-09-147-236-1 (1-16836)
QY 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThr---ProGluGln 65
Db 8324 ACCCAGTGGGACCAAGCGCGCGGGGCTGCCCGGGATCGTTGGCTCCAAACCCGAGAA 8383
QY 66 IleGlnAlaArgLeu-----AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGln 82
Db 8384 TACCGCGCCAGCTCGCTTTGGCCAGCGCCCTTTGACCTATATAATACCTCGACCCCGATGAA 8443
```

D	b	9353	GATGTGGCTGAACCGGCGCGTGATG-----CAGCCCATCTTGGCCAATCTGTCAACC	9400
Q	y	401	AlaTyrAsnLeuGlnAla-----ValArgAlaLeuSerAsnLeuIleAla	416
D	b	9407	GCGCAGGACCGCAGCGCGGTATCTGTATACCTATGTCAGTGGCAATGATGCATGACC	9466
Q	y	417	ThrArgTyrPheAsnMetValAsnThrGlulIleValPheProGluArgGlulIleGln	436
D	b	9467	CGCCAGCTTTCGCCGCGTCTGTGCGCGCGGATTATCCCCTCCGGATCGCTGTTCCATGTGCC	9526
Q	y	437	AsnAspGlnValSerPheGluGlnSerSerSerArgTThrGluProAlaGlnValAsp	456
D	b	9527	--GAGAAATGGAATCAAGCAGGACCTGGCCACGCCCTGTCGATGGTATCCAACCCG	9583
Q	y	457	GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetasp	476
D	b	9584	GTCCCGCTGATCGCAGGCGCTTTCCACAGCTGATCCGACCGCGCGCTGGCGTGGCG	9643
Q	y	477	IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValala	496
D	b	9644	GTGGCCGACTGTTCGCCCAGCGTGGCAGACATGGTGCATGCGCGCATGCCCTCGGTATC	9703
Q	y	497	AlaLysAlaArgHisLeuTyrAspMetProAspAspArgVal-----	510
D	b	9704	GCCTCAGCGCGCACCATCGACCTTTGCGCCGACACGCGCTGTCTACGCCACCGAATAC	9763
Q	y	511	LeuAlaIleAsnHisAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAla	530
D	b	9764	ATGAAGATCAGCAACCCGCTGGCAGCGCGCACGC---CTGCTGGCCCCCTGGGGATGGC	9820
Q	y	531	Val---SerAlaValAlaArgAlaIleLeuProAspGluSerGlu-----	544
D	b	9821	ACGGTTTCGGCGCGGCAATCGCTGTGCTGCCGAGCAGATGCAGACATTCGACCAACTG	9880
Q	y	545	AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAsp	564
D	b	9881	CGCATGGGCATCTCGGTGGCGCAGTCCGATCTGCTCAACACGCTGGCGATCAGCGCAG	9940
Q	y	565	ValTyrGln-----	567
D	b	9941	GCCTACGATCATCTGCGCGCGCCCTGTCAGCGCGACCCGAGGCGCACATCCGCCAAGCTG	1000
Q	y	568	-----SerLysLysValProLeuTyrVal	575
D	b	10001	GCGTCTCGCGCGCTGTATAACGCCACGCGCAAGCGCGGCAAGCGCTCGAGATCGACCTT	1006
Q	y	576	PheValAlaSerAspLysProArgspGlyGlnIleGlyLeuGlyTrpGlySerAspThr	595
D	b	10061	GCGGTCTGCGCCACAACCGCGAGNATCTTGATGC- GCGGACGCGCGCGTGCAGCGCGC	1011
Q	y	596	GlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla	615
D	b	10120	GGTCAACAGCAACCAACAGCTTTGCCACCG-----TCTCGCATGGATGGCGTGCA	1017
Q	y	616	Gly-Ala-----	617
D	b	10174	GGAAGCCCGATGGATGCCCGCGCTGCGTGGCCATGCGCTGACCAAGCGCGATGG	1023
Q	y	618	-----GluLeuArgLeuSerGluAspLysGlyValLysLe	630
D	b	10234	CCATGGTCACGCGCACCATCGAGGACCTGGCGCGCTATGACCTGCGCGCTGCAGCAGGT	1029
Q	y	630	uTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTy	650
D	b	10294	CGAGGCGACGCGGCGCGCTGTGGTCCGTCGGG-----GCACA	1033
Q	y	650	rGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuG1	670
D	b	10333	TGAAGAAGCCCTTGC CGCGCATCGACCAACCCGTTC-----CAGTCCGCTGGCTACCG	1038
Q	y	670	uHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTyrSerLeuAr	690

Db 10387 GCATCAGGTGGAACCTGGCGCGCGGTGACCGGTGGC-----TCCTACAGTGC CGA 10437  
Qy 690 gTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGlnAspLeuProVa 710  
Db 10438 G-----GCGGCATCGCCCGATAG----- 10456  
Qy 710 lAspPheValAsnGlyLysProSerGlnAlaLeuLeuAlaGlyValAla-----Va 728  
Db 10457 -----TCGGACACAGATGCTCTCCCATTTGCTGGCCAGAT 10491  
Qy 728 lHisLysThrValAlaAspAsnLeu-----ValAsnProMetArgGlyTyrArgGl 745  
Db 10492 CCAC---ACGCTGGCTGAACACCTTGACCCCTCCATTCATGGGCTGGGCTTCCGGTC 10548  
Qy 745 nArgTyrSerLeuGluValGlySerGlyLeuValSerAspAlaAsnMetAlaIleAl 765  
Db 10549 GCGT---TCGGCGGAGCATGGC---ATGGCGCGCTGACGGAGCAACATTCCTCCATCGT 10602  
Qy 765 a-----ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTy 780  
Db 10603 GGGCGCGCTGCGCTGCAGCGCGGTGCTTCGGCC----- 10636  
Qy 780 rGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAs 800  
Db 10637 -----CTGACCTTCGATCATACGCCACCATGATGCTGTCGGG 10674  
Qy 800 pAsnPheAsn-----HisValProTyrArgLeuArgPhePheAlaGlyGl 815  
Db 10675 CCAGCTCAACACAGCGCTCCGTCTATGATGTCGG---CGTTATGGCACGTTTCATGGCAAC 10731  
Qy 815 yAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTy 835  
Db 10732 GCAGGCTGCCAACCAACAGTCGCGGGCCACAGTTCGTGTGGCGGCTTGAT-----TT 10782  
Qy 835 rLeuThrGlyGly-----GlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPh 853  
Db 10783 CTGAGCGCCCAACCATACCCAGCGCATCGCGGCTGGTGCAGGCGAG----- 10828  
Qy 853 eMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPh 873  
Db 10829 -----GCCGGGTT 10836  
Qy 873 eThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrp----- 887  
Db 10837 TCGCGCGGATGTGCAGTTCGGCAATAGC-----TCGGTGGCGGTGATGTCTG 10884  
Qy 888 -AlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys 903  
Db 10885 CGCCTCGCCCATCGGCTTCCCATTTACCAACAGTGTGGGCGGGGTCGAG 10933

Search completed: May 6, 2003, 03:17:43  
Job time : 5587 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 6, 2003, 01:42:26 ; Search time 179 Seconds  
(without alignments)  
6066.535 Million cell updates/sec

Title: us-09-914-168-2

Perfect score: 4727

Sequence: 1 MSKPVLFPANRSMFVALAAY.....TGVKEGNPKLHFHFIQTGF 919

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09914168/runat\_28042003\_151617\_6886/app\_query.fasta\_1.1095  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Diosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09914168\_@cgn\_1\_1\_99\_@runat\_28042003\_151617\_6886  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length DB	ID	Description
1	238.5	5.0	6617	9	US-10-210-296-1
2	140	3.0	17310	7	US-08-781-986A-23
3	139	2.9	5245	9	US-10-080-505-12
4	137.5	2.9	7104	10	US-09-815-242-4580
					Sequence 4580, Ap

5	137.5	2.9	7107	10	US-09-815-242-8291	Sequence 8291, Ap
6	136.5	2.9	38584	9	US-10-114-170-50	Sequence 50, Appl
7	136	2.9	8673	10	US-09-815-242-7474	Sequence 7474, Ap
8	136	2.9	12404	9	US-10-108-605-236	Sequence 236, App
9	135	2.9	5629	9	US-10-114-170-243	Sequence 243, App
10	135	2.9	6228	10	US-09-815-242-4760	Sequence 4760, Ap
11	135	2.9	6561	10	US-09-815-242-8815	Sequence 8815, Ap
12	134	2.8	3121	10	US-09-765-272-117	Sequence 117, App
13	133.5	2.8	14536	9	US-10-160-758-6	Sequence 6, Appli
14	133.5	2.8	14536	9	US-10-160-758-7	Sequence 7, Appli
15	133	2.8	17388	10	US-09-815-242-8512	Sequence 8512, Ap
16	132.5	2.8	1488	10	US-09-815-242-9166	Sequence 9166, Ap
17	132.5	2.8	2571	10	US-09-815-242-7037	Sequence 7037, Ap
18	132	2.8	7434	10	US-09-815-242-4761	Sequence 4761, Ap
19	132	2.8	7437	10	US-09-815-242-8869	Sequence 8869, Ap
20	131.5	2.8	6077	10	US-09-815-242-4648	Sequence 4648, Ap
21	131.5	2.8	9477	10	US-09-815-242-8513	Sequence 8513, Ap
22	131.5	2.8	31096	7	US-08-781-986A-59	Sequence 59, Appl
23	131	2.8	7302	10	US-09-815-242-4780	Sequence 4780, Ap
24	131	2.8	18846	10	US-09-815-242-8898	Sequence 8898, Ap
25	130.5	2.8	3552	10	US-09-815-242-4174	Sequence 4174, Ap
26	130.5	2.8	3567	10	US-09-815-242-8027	Sequence 8027, Ap
27	130.5	2.8	7311	10	US-09-815-242-4779	Sequence 4779, Ap
28	128.5	2.7	3336	10	US-09-954-456-1807	Sequence 1807, Ap
29	127.5	2.7	5832	10	US-09-881-752A-307	Sequence 307, App
30	127.5	2.7	8673	10	US-09-815-242-7313	Sequence 7313, Ap
31	127	2.7	2545	10	US-09-974-300-1524	Sequence 1524, Ap
32	126.5	2.7	11185	10	US-09-880-107-3311	Sequence 3311, Ap
33	124.5	2.6	9534	9	US-09-954-531-1380	Sequence 1380, Ap
34	124.5	2.6	10302	10	US-09-782-378A-23	Sequence 23, Appl
35	124.5	2.6	14155	9	US-10-108-605-102	Sequence 102, App
36	124.5	2.6	50937	9	US-09-808-880-1	Sequence 1, Appli
37	124	2.6	1815	9	US-10-107-649-1	Sequence 1, Appli
38	124	2.6	5346	10	US-09-738-877-2	Sequence 2, Appli
39	124	2.6	6608	10	US-09-738-877-1	Sequence 1, Appli
40	124	2.6	6608	10	US-09-880-107-3439	Sequence 2, Appli
41	123.5	2.6	3349	9	US-09-952-267-2	Sequence 363, App
42	123.5	2.6	3753	10	US-09-801-368-363	Sequence 48, Appl
43	122.5	2.6	2609	10	US-09-789-561-48	Sequence 1, Appli
44	122.5	2.6	6026	9	US-09-995-749A-1	Sequence 3499, Ap
45	121.5	2.6	2556	9	US-09-738-626-3499	

#### ALIGNMENTS

RESULT 1  
US-10-210-296-1  
; Sequence 1, Application US/10210296  
; Publication No. US20030021802A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Products Inc.  
; TITLE OF INVENTION: LAWSONIA INTRACELLULARIS PROTEINS, AND RELATED  
; TITLE OF INVENTION: METHODS  
; TITLE OF INVENTION: AND MATERIALS  
; FILE REFERENCE: PC10589A  
; CURRENT APPLICATION NUMBER: US/10/210,296  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US/09/689,065  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 6617  
; TYPE: DNA  
; ORGANISM: Lawsonia intracellularis  
US-10-210-296-1

Alignment Scores:  
Pred. No.: 1.86e-13 Length: 6617  
Score: 238.50 Matches: 198  
Percent Similarity: 34.72% Conservative: 137  
Best Local Similarity: 20.52% Mismatches: 391  
Query Match: 5.05% Indels: 239

DB: 9 Gaps: 38  
US-09-914-168-2 (1-919) x US-10-210-296-1 (1-6617)  
QY 50 AsnGlnAlaLysAlaGlyAsnProValLeuLeuThrProGluGlnIleGlnAlaArg 69  
DB 3764 AATGCTGCTCAAGAGGATCCTTCTATTGTGGTCTCCCATTTCAAAATTAATGGCTCA 3823  
QY 70 LeuAsnAlaAlaGlyLeuAsnAlaLys---ProGlnSerGlnAlaLeuAspValValAsn 88  
DB 3824 TCAAAATGATGAAGAGTTACAAACAGAACTACCAATGCTTCTTGCAACTGCATTAAGAAT 3883  
QY 89 PheAspGlnSerProIleSerArgIleGlyGluGlnSerProLeuGlyLeuAsp 108  
DB 3884 -----AAGGGA 3889  
QY 109 MetSerValIleGluThrThrProLeuSerLeuGluLeuPheAlaGlnGluSer 128  
DB 3890 TTTTCGTGTCATCCCTAATAAATCTGCATTAATCTT-----CTATATAACAAATAATC 3943  
QY 129 ThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn 148  
DB 3944 TCCCAACTTAATATTCTACTGCAGGAGGTA-----GCTCAACAACCTCCAT 3991  
QY 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168  
DB 3992 GCTGACTATGAGTA----- 4006  
QY 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188  
DB 4007 TACGGCATTTCATCAACACAGGTGAAT-----TTTAGT 4042  
QY 189 GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSer-----HisGlnLys 205  
DB 4043 ATTGATAGTAGGCTTATTGTAGTACAGGTGTAGCATCTGCACGTCCATTAATACATAGAA 4102  
QY 206 ThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGluSerAlaMet 225  
DB 4103 AAACCAAAATTAATAGCTAAATATTGCTGTACACAGAACTGTCTGAACGTATTAAGT--- 4159  
QY 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaVal 245  
DB 4160 -----AATGGCCTTAAAGAAA---AACACTATTGCTGATGATACGTATTCATGGCCTT 4210  
QY 246 GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle 265  
DB 4211 AAAGTTCTTGATCCTGATGAATCTTACACGACTACTATTATAAAGGGAGATCATACT 4270  
QY 266 IleHis-----AspLeuGlyGluProValTyr 274  
DB 4271 GATCATGCCAAATTAATGCAGAAATCAAAAATATATGGGAATTAGGA-----TATTTT 4324  
QY 275 IleAspTyrArgAla---ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThr 293  
DB 4325 AGTGATGCTCTGCAAGTATTGAAGAAAGCGGGAAGGA-----CGATTACTTGTGA 4375  
QY 294 ThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGlu 313  
DB 4376 TTTACTGTACAAGAAAGCCCTCAAAATTACAGATGTTTGTGTTCAAGGCTCAAAAGCTGTA 4435  
QY 314 ThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyr 333  
DB 4436 AGTATCCATCAATCTTGTGCTCAATCAGTTCTAAAAAGGA----- 4477  
QY 334 LeuAspArgSerVal-----AspValIleLeuProAspAsnThrAlaAspValSerLeu 351  
DB 4478 -----TCAGTTATTAGTAGACTATTGTCACAGATATATCAAAAAAATTAACGCAC 4528  
QY 352 IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspPro 371  
DB 4529 CTCTATAGAAAGAGGCTACTATCTCGCTGAA-----GTTAATTATGAATAAAGAG 4582  
QY 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlu 391

DB 4583 AAAGAAAATACTTCTTCTGCA----- 4606  
QY 392 GlnLeuLeuThrValAsnMetGlyClyAlaLysAlaLysLeuGlnAlaValArgAlaLeuSer 411  
DB 4607 CTATTGTTAAACAGTAATGAAGGAAATACTTTATATTAAGATGTCGCGATTGAAGGA 4666  
QY 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431  
DB 4667 CTTCAACAATAAAGAGCTAAA-----ACTTTAAAAAAGAGCTTAGCATTAAACAGAA 4717  
QY 432 Arg-----GluGlnIleGlnAsn 437  
DB 4718 CGTAATTTTATCATGTTTACTGGAACAGGTGTATTACGTGAAGATAATCTTGAACGT 4777  
QY 438 AspGlnValSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspGlu 457  
DB 4778 GACTCTATAGCAATC-----TCTGCCCTATGCCATGAATCATGGCTATGTAGATATT 4828  
QY 458 SerThrLeuGluPro-----ValIleGluThrValIleGluLeu 469  
DB 4829 CAAGTTGCTTCCACCTGAAGTAACATTCAATGAAAGAAAGAAATGTTATTACATTTAGAGTA 4888  
QY 470 ThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln 489  
DB 4889 AAAGAAGGTAAAGCCCTATAAAATAGAAAAATAGACTTTTAAAGGAGATCTTTATTTCAGACA 4948  
QY 490 AspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArg 509  
DB 4949 AATGAACAACCTCTTAAAGTAACAAAAATTCATGATCATATAAAACTATGACCACTATTTT 5008  
QY 510 ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAsp 529  
DB 5009 TCTCTTCTGTATGCAAGATGAT-----GTAAGACCATTAACAGATTTTATTATTCAGAT 5062  
QY 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu----- 546  
DB 5063 TATGGTTATGCATTTGCTGAAGTAGATCTT---CAAAACACCAAAATGAAGAAGATGCA 5119  
QY 547 -----ValIleAspLeuProGluArgThrAlaLeu----- 556  
DB 5120 ACAATTGATGTTACTTCTTATTGATAAAAACAAGAGTCTTCTTCGTAGATAATTT 5179  
QY 557 -----AlaAsnArgLysThrProAlaAspValTyr----- 566  
DB 5180 GTTGAAGGAAATACTCGTACTAGAGATAATGTTATCCTCGTGAATACGCGCTTGCCTGAT 5239  
QY 566 ----- 566  
DB 5240 GGAGATCTTTTAAATGTCACATCTCCGAGGCTCTAATGAATCCCTTAACCGCCTTGGC 5299  
QY 567 -----GlnSerLysLysValProLeu 573  
DB 5300 TATTTTAAACAGTAGATACAGATACACTGCCTACAGGAAAGATGATGAAGTTGATCTA 5359  
QY 574 TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySer 593  
DB 5360 CTTGTAAAGTT-----CAAGAGCTCGAACAGGTGCAATCACAGGTGGTGTGGTTAC 5413  
QY 594 AspThrGlyThrArgLeuValThrLys-----PheGluHisAsnLeuIleAsnArg 610  
DB 5414 TCAACACATTTAAATTTGGTGTTCAGGAAGTATCTCAGAAAGAAACCTATATCGGGGAAA 5473  
QY 611 AspGlyTyrGlnAlaGlyAlaLeuArgLeuSerGluAspLysLysGlyValIleLysLeu 630  
DB 5474 ---GGTTATATTTAAAGTATTGAAGGTTTATTCTTAGTAAGTCATCTCTCTTGTATCTT 5530  
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650  
DB 5531 TCTTTTACCCTCTCGTGTATGATACAGAC----- 5563  
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeu----- 664

```
Db 5564 -----TTTGGCTTTAGTAATAACATTTATACGCTACGAGATGAATGGGATGAC 5611
Qy 665 -----SerThrArgThrLeuGluHisGluIleSerArgSerIle 677
Db 5612 TTCGTAATAAAACCTTTAGGAGATACCATAGCTCTATTTCACCTATATA----- 5659
Qy 678 IleGlnAsnGlyTyrPAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLys 697
Db 5660 -----GCAGATAATTCATCTATCTTTGCTGCTATCGAATGATCAATATCGT 5707
Qy 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro 717
Db 5708 CTATATGATATTCATCTACAGCACCAGCGCTTCTATCTTGACTAT---CAAGGGAAAAAT 5764
Qy 718 SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 5765 ATTTCTAGTAGTAAGTGGTGTCTTTACTTTTGTATCTACA-----GACAGTCGTGAG 5818
Qy 738 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal 757
Db 5819 AGACCATCTAAAGGCATATTGCAAACTAATTCTTGAATATGAGAGTGGTGGTCTTGGT 5878
Qy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 5879 GGTAATGATAACTTCTTCAAGCCCAATTGCTCAACTACAGGATTTTACTCAATT----- 5932
Qy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIle 797
Db 5933 -----TCAGAGTAAACCATATATATACATTTGCGCTACAGGTCGACGCT 5983
Qy 798 TrpSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGln 817
Db 5984 TATAAGATAGTAAACCCCTGTCGCATATTGACCCATTTTATTGGTGGTATAGAT 6043
Qy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys---GlyTyrLeu 836
Db 6044 AGTATTAGAGATATGATACAGAGATCTTGCCACCAAGATCCTCGCTTTGGAGATGAA 6103
Qy 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp 856
Db 6104 ATTGCTGTGATAGATGGCTTTTCTTAACCTAGAGTATATTGGACATTTCCAGCCAGAG 6163
Qy 857 LeuArgLeuAlaVal-----PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThr 874
Db 6164 CTAGTCTGTGATAGTTCCATTCTATGACATAGGATTCACCAAGATTCCTGTACAAACT 6223
Qy 875 AsnAsp-----ThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerPro 890
Db 6224 TCTAACCCATTCTCTAAACTCAACCAATCATATGSCCTTGAACCTCGCTGGCTTCACCA 6283
Qy 891 ValGlyGlnValArg 895
Db 6284 ATGGGAGATTTGCGA 6298
```

## RESULT 2

```
US-08-781-986A-23/c
; Sequence 23, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
```

```
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/781,986A
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-23
Alignment Scores:
Pred. No.: 0.00607 Length: 17310
Score: 140.00 Matches: 177
Percent Similarity: 32.16% Conservative: 124
Best Local Similarity: 18.91% Mismatches: 276
Query Match: 2.96% Indels: 359
DB: 7 Gaps: 45
```

US-09-914-168-2 (1-919) x US-08-781-986A-23 (1-17310)

```
Qy 205 LysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla 224
Db 12111 AAAAGTAAATAAATAACAACTTATAAAGAACAAATTCGCGTATTTTTCACACATATGGCT 12052
Qy 225 MetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAla 244
Db 12051 CAAAGTGCAGAT-----AATCAAGCGCAGTGCACAAACATATGGCG 12010
Qy 245 ValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspVal 264
Db 12009 CAAATTATGCTGACATAGCATCATATGATCATATCAGCGCA----- 11968
Qy 265 IleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValAlaGluValArgGly 284
Db 11967 -----CGTTATTGTCAAAGTAACACAATAGACTAC----- 11938
Qy 285 GluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGly 304
Db 11937 -----GTAGAGGATCGT-----TATATCGTTTCAGAACAACTCTTTATACTTAGGT 11893
Qy 305 AspValPheHisHisGlyLysTyrGluThrLysLysAsnLeu----- 318
Db 11892 CATCCATTTTCAT-----CCGACTCCTAAGAGTCAAGTGGGTTTTCAGAAGCA 11845
Qy 319 ---IleGluAsnAlaSerAlaGlu---HisGlyTyrPheAspGlyArgTrpLeuAspArg 336
Db 11844 GATTTAGAGAAATATGCACCGCAATGTCATACATCATTCCAATTGCAATTTAGTGTG 11785
Qy 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356
Db 11784 CATCAAGATGTTCTGCTCAGCGC-----TATGTAGAAGGT 11749
Qy 357 ThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGln--- 375
Db 11748 AAGAAGATCAGGTTGAGAAAGTGTGTCAATTTAGCAGACATATATATCAGAGATA 11689
Qy 376 -----LeuThrThrAspProAspLysLeuProValLysArg----- 387
Db 11688 CCCAAGATTTTATTTTATTATACCAACACATCTCTATCAANTCAATCTGTTCGACAGCAT 11629
```

QY 388 -----GluLeuLeuGluGlnLeuLeuThrValAsnMetGly----- 399  
Db 11628 CCACAGTATATGCAATATAGTGAACAAGGTTTAAATAAAGACCTTGGCGTGAT 11569  
QY 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeu----- 410  
Db 11568 TCAGTGTACCGCGCTTTCGGTTAGAACTGATTTTCAAAAGCATTAACACATTTATTTA 11509  
QY 411 -----SerAsnAspLeu----- 414  
Db 11508 AAATTACCGATACACGTTAAAATCACTAAATTTTATACGTACGAATGACCTTGAACAGATT 11449  
QY 415 -----IleAlaThrArg 418  
Db 11448 GAAGGACAATTTGATCGCCGCGAAGTTATCGCATCAGTCAAAGATGAGGTTGAAACACCC 11389  
QY 419 TyrPheAsnMetValAsnThrGlu----- 426  
Db 11388 CATTTTAAATTTGATGTTTGAAGAAGGATATCGTGCAATTTGTCGGAATCCATTAGGGCAA 11329  
QY 427 IleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer--- 445  
Db 11328 ACAGTTGAACCTGAATGATGATTTTAAACAATAGTGCCATGATGTTCTGCGTGAAGGATA 11269  
QY 445 ----- 445  
Db 11268 CGAATTACCATGCTGATAAAGATATTCATGATGTCGCTCATTTATTTGAAACGATGCCT 11209  
QY 446 SerSerSerArgThrGluProAlaGlnValAspGluSer--- 458  
Db 11208 GATTCACCGATGCTCAAGTTATCACAAAGTATGAGCAAAAGTGGTTTAGCACCAGAAGCA 11149  
QY 459 -----ThrLeuGluProValIleGlu----- 465  
Db 11148 TGGCTTGAATGTTATTTGAATCGTACATTTTCCCGCATTTAAAGCTGTTAGTAAACACA 11089  
QY 466 -----ThrValGluLeuThrAspGlyIle--- 473  
Db 11088 GGCATTAGTCTAGAAGCACATGTACAAAATACATTAATTTGAATTAAGAGATGCGATACCC 11029  
QY 474 -----LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln 489  
Db 11028 GACGTATGCTTGTGAGAGATCTTGAAGCATTTGTCTATCTAGACGATTTGCTACTGAA 10969  
QY 490 AspLysLeu-----AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp 507  
Db 10968 AAACAGCTTGTGCGAAATGTTGGCGAGCATCAAGCCCTGTTGTATAT----- 10921  
QY 508 AspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg-----SerIle 523  
Db 10920 -----GCACATGATGAAGCATGCGTCGTCTAAATATTACGTTGTA 10879  
QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543  
Db 10878 GTAAATCACTTAGCAGATTTAGTATCACTATTGGTAAAGCG-----ACT 10834  
QY 544 GluAsnGluValIle-----AspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561  
Db 10833 AGAAATGAAGTTGTTATGGCAACTTGTAGGCGATCTTGTATGACTTGGAAAAAAGAA 10774  
QY 562 ProAla-----AspValTyrGlnSerLysLysVal 571  
Db 10773 TAGCGGAATAACGCGAGTATTTGTTGACTGTGTAGAAGATTTATATCAACGCGGACCAT 10714  
QY 572 -----ProLeuTyrVal 575  
Db 10713 GCGGCTAAACGCAATTTGATGAGTAATTAATGATGTTGGTGCAACCCCTATTATTATACA 10654  
QY 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGly----- 592  
Db 10653 CATATACCAATCCAAATTTGTCTATACAAAGGAGGTATCTGTTGTGAATCAAAACAAATTC 10594  
QY 593 -----SerAspThrGlyThrArgLeuValThrLys-PheGluHisAs 606

Db 10593 TAATCGTGTAAAGACTAGAGTGTACCACTGTCATCATCTATTATTATGAGAATAT 10534  
QY 606 nLeuIleAsnArgAspGlyTyrGlnAlaGly----- 616  
Db 10533 TGTGTGTATAAAGCGTCATATCAAGACGGTGTGCGTCTATTTCACATAGAAGCACATGA 10474  
QY 617 -AlaGluLeuArgLeuSerGluAspLysLysGlyVal-----LysLeuTyrAlaThrIly 634  
Db 10473 TTCAGAGTATCGTTTACTGCTGAAAGACACATAGCTTTGATCGTATAGTATACATC 10414  
QY 634 sProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln----- 652  
Db 10413 ACCAATTGAGCGTGTCTAGGAGATGAGGAGATACAAACACAGACTATACAAATATT 10354  
QY 653 -----GluValPheGlyHisSerThrAsnGlyPheAspLeuSer-----ThrAr 667  
Db 10353 GAGAGAGGTTGATTTTACATTTCTCTAAAATGATGAAACATAGAAACAATTTATTTCGA 10294  
QY 667 gThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTy 687  
Db 10293 GTTATTACAGACAGATTAT-----AAAGATACGCA 10264  
QY 687 rSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAs 707  
Db 10263 AAGTATGCAGTATCGA-----GAATCAAAACCCACGAGCAACACCTGAGACATTTAACGA 10210  
QY 707 pLeuProVal----- 710  
Db 10209 CTATGAATTTTATGCCATGCGAAGGCATCAGTATCATCTCAAGTTACAAATACGCTTAGG 10150  
QY 711 -----AspPheValAsnGlyLysProSerGI 719  
Db 10149 ATTTACGTTGAGTGATAATTTGAAATTTGCTCTGTATTTGTACCAACCGTTAACTGCA 10090  
QY 719 nGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsn---LeuValAs 738  
Db 10089 GTGGTTAGTATCGCAAAAGATAAAGTAAAGCAGCGGTATCAAGAAATGTTGTAGTTAA 10030  
QY 738 nProMet-----ArgGlyTyrArgGlnArgTyrSerLeuGluVal 751  
Db 10029 CGAAATGTTACGTCACAAAGTTGGCGCATTAAGACTTAT---GAACATTTTGTACAGCAAT 9973  
QY 751 lGlySerSerGlyLeu---ValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSe 770  
Db 9972 TGAAGCATCTGCGAAACATGTAATGATGTTGAGATGATACCTGTACACCCA----- 9921  
QY 770 rGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGI 790  
Db 9920 -----TGGCAGTTT---GAACATGTCATCCAAAGTTGATTTGGCTGAAAGAAAGCGTTAA 9871  
QY 790 yGlyIleGlnAlaGlyTyrIleTrp-----SerAspAsnPheAsnHisValProTy 807  
Db 9870 TGGCACA-----GTACTATGTTAGGGAAAGTGTAGCTATATATC----- 9828  
QY 807 rArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLe 827  
Db 9827 -----CCTCAACAATCGATTCTG-----ACGAT 9805  
QY 827 uSerProIleSer---AspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThr 846  
Db 9804 GTCGCAATAGACACGACAAA-----ATATTATTAAAGGTACC 9766  
QY 847 AlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGly 866  
Db 9765 AATAAGTATACGAACACTTC----- 9745  
QY 867 AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArg 886  
Db 9744 -----AACGAAACGAGTGTGGCCCTCATACAAATTCGAAA 9709  
QY 887 TrpAlaSerProValGlyGlnValArgValAspValAlaThr 900

Db 9708 TGCAGCGCAAAATTACGGATTGGTTAAAGCAGATACAGCAACA 9667

## RESULT 3

US-10-080-505-12  
; Sequence 12. Application US/10080505  
; Publication No. US20030073166A1  
; GENERAL INFORMATION:  
; APPLICANT: St. Gene, Joseph W.  
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS  
; FILE REFERENCE: A-5941-1/RFT/DCF/DHR  
; CURRENT APPLICATION NUMBER: US/10/080.505  
; CURRENT FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: US 08/296,791  
; PRIOR FILING DATE: 1994-10-25  
; PRIOR APPLICATION NUMBER: US 09/839,996  
; PRIOR FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 5245  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (430)..(4740)  
; OTHER INFORMATION:  
US-10-080-505-12

Alignment Scores:  
Pred. No.: 0.00133 Length: 5245  
Score: 139.00 Matches: 157  
Percent Similarity: 34.10% Conservative: 153  
Best Local Similarity: 17.27% Mismatches: 352  
Query Match: 2.94% Indels: 248  
DB: 9 Gaps: 36

US-09-914-168-2 (1-919) x US-10-080-505-12 (1-5245)

Qy 13 MetProValAlaLeuAlaAlaTyrLeuProLeuMetThrSerGlnAlaLeuAlaGlnGln 32  
Db 2384 ATGAGATTGGATCACCAGCATTTAAAGCTGAAAGCTCCAAATTAAGCGGGAAGTG 2443  
Qy 33 AsnAsnProAlaAsnIleAsnHisValProAlaHisAspThrAlaIleAsnGlnAla 52  
Db 2444 CGGTG-CTTCTCGCAATGTTCTTCAATTGAGGGAATTCGACAGTCAGCAATAATGCA 2502  
Qy 53 LysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAla 72  
Db 2503 AATCCCGCATTTGGTGTGTCGCAAAATCAGCAAAATACCAATTCGACCGTTCAGATTGG 2562  
Qy 73 AlaGlyLeu-----AsnAlaLys 78  
Db 2563 ACAGGATTAAGCACTTGTAAGACTGTGGATTAAACCGATACAAAGATTATTATTCATA 2622  
Qy 79 ProGlnSerGlnAlaLeuAspValValAsnPheAspAspGlnSerProIleSerArgIle 98  
Db 2623 CCGCAACACAAATTAATGGCTCTATTAAATTAATGATAATGCAACAGTGAATATTTCAT 2682  
Qy 99 GlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGlu-----113  
Db 2683 GGT-----TTAGCAAACTTAATGGTAATGTCATCTTAATAAATCATPAGCAATTTACA 2736  
Qy 114 -----GluThrThrProLeuSerLeuGlu-----121  
Db 2737 TTGAGCAACATGCCACCAACAGGCAATATCCAACITTCAAATCAGGCAAAATGCAACG 2796  
Qy 122 -----GluLeuPheAlaGlnGluSerThrGluMetGlyIle 133  
Db 2797 GTGGACAATGCAAAATTTGAACGGTAATGTGCATTTAAACGGATTCTGCTCAATTTCTTTA 2856  
Qy 134 AsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValVal 153  
Db 2857 AAAAACACCAATTTTTCGCAACCAATTCAGGGCGACAAAGACAAACAGTAGTCACGTTGGAA 2916

Qy 154 ProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArg-----171  
Db 2917 AATCGACTTGGACAATGCCTAGCGATCCACATTCGAGATTTAACGCTAAATAATAGT 2976  
Qy 172 -----LeuPheAsnAspGlyValAsnLysValProArgLeuLysAla 185  
Db 2977 ACTGTTAGCTTAAATTCAGCTTATTAGCTAGCTCAAAATAATGCGCCACGT-----3027  
Qy 186 LysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys 205  
Db 3028 -----CACCGCGTTCATTAGACAGCGGAACACCGCAACATCGCGCAACAT-----3075  
Qy 206 ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMet 225  
Db 3076 -----CGTTTCAACACATTG 3090  
Qy 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaVal 245  
Db 3091 ACAGTAATGTTAAATTGAGCGGG---CAAGGCACATTCCAATTTACTTCTTTATTT 3147  
Qy 246 GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle 265  
Db 3148 GCCTATAAAGCGATAAATTAAATTTATCAATGCGCTGAGGCGGATACACATTATCT 3207  
Qy 266 IleHisAspLeuGly---GluProValTyrIleAspTyrArgAlaValGluValArgGly 284  
Db 3208 GTTCGCAACACAGCAAGAACCCGAGCCCTTGAGCAATTAACCTTTGGTT-----3258  
Qy 285 GluGlyAlaAspAspLys-----AlaPheThrThrValAlaAspGlu 298  
Db 3259 GAAAGCAAAAGATAAACCCTTTATCAGACAACTCAAAATTTACTTTAGAAATGACCAC 3318  
Qy 299 Val-----ProLeuLeuIleGlyAspValPheHisHisGlyLysTyr-----312  
Db 3319 GTTGATGCGAGTGCATTACGTTATATAATTAGTAAGATAATATGCGCAATTCGCTTGCAT 3378  
Qy 313 -----GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326  
Db 3379 AACCCAATAAAGACGAGGAATTCGCAATGTTAGTAAGACGAGACGACAGCAAAA---3435  
Qy 327 GlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThr 346  
Db 3436 -----CGAACATTAGAACGCCAAACAAAGTTGAACAGACT 3468  
Qy 347 AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe 366  
Db 3469 GCTGAACACAAA-----ACAAGTAATGCAAGAGTGGGTCAAAAAGACCGGTGTT 3519  
Qy 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386  
Db 3520 TCTGATACCTCGCTGATCAAAAGCCAGTTAGAGTA-----TTACAA 3561  
Qy 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406  
Db 3562 GCCCAACAAAGTTGAACCGAGCTGCTGAAAAACAAAAATAAGGCAAAAAAGGCGGTCA 3621  
Qy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArg---TyrPheAsnMetValAsnThr 425  
Db 3622 AAAAGACGGGTGTTTTCGATACCTCGCTGATCAAAAGCCAGTTAGACGATATTCAAGCC 3681  
Qy 426 GluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer 445  
Db 3682 GAACAAAGTTGAACCGAGCTGCTGAAAAACAAAAATAAGGCAAAAAAGTGGCGTCAAAA 3741  
Qy 446 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465  
Db 3742 AGAGCCGCGAGA-----3753  
Qy 466 ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer 485  
Db 3754 -----GAGTTTTCGAT-----ACCCCGCTTGCATGACTAAGCCGGTTA 3789

Qy	486	AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet	505
Db	3790	AAAGTATTAGAAGTCAAACTTGGAGTTTAAATCCCAA---CAGCAAGTGCAAAAAGAA	3846
Qy	506	ProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg	522
Db	3847	CCTCAGATCAAGACAAACACGCAAAACAAAGACTTGATCAGCGCTTATTCAAATAGT	3906
Qy	523	IleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlu	542
Db	3907	CGCTTATCAGAATTATCTCCAACAGATAAATAGTAGTGTCT---GTTCAAGATGAA	3960
Qy	543	SerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrPro	562
Db	3961	TTAGATCGCTTTTGTAGATCAACCACTCTCCGCTGGCACAAT-----	4008
Qy	563	AlaAspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLysPro	582
Db	4009	-----ATCGCACAGGATAAAGA	4026
Qy	583	Arg---AspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThr	601
Db	4027	CGGTATGATTCTGATCGGTTCGCTGCTTATCAGCAGAAACGAACCTTACGT-----	4077
Qy	602	LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu	621
Db	4078	-----CAAAATGGGGTGCAAAAAGCGCTTA	4101
Qy	622	SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn	641
Db	4102	GCTAATGAGCAAAATGGGCA---GTTTCTCGCATACGCCCTCAGATAATACTTTTGT	4158
Qy	642	AspGlnLeuArg---AlaThrLeu-----GlyTyrGlnGlnGluValPhe	655
Db	4159	GAACAGGTTAAATCAGCGCATTAACGATGATGTCGGGTTTGCCCAATATCAATGG	4218
Qy	656	GlyHisSerThrAsnGlyPheAsp-----LeuSerThrArgThrLeuGluHis	671
Db	4219	GGCGATTTCAAATTTGGTGTAAACGTGGGAACGGGAATCAGTCGCGAGTAAATGGCTGAA	4278
Qy	672	GluIleSerArgSerIleIleGln-----AsnGlyGlyTyrPasnArgThrTyrSer	688
Db	4279	GNACAAACCGCAAAATTCATCGAAACGATAAATATGCGGTGAATGCAAGTTATCAG	4338
Qy	689	LeuArg-----	690
Db	4339	TTCCGTTTAGGCAATTTGGCATTCAGCCTATTATTTGGAGTAACTGCCTATTTATTGAA	4398
Qy	691	-----TyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTyrGlnAsp	707
Db	4399	CGTGAAATATCAATCTGAGGAAGTGAAGTGAACACGCTAGCCCTGCATTTAATCGC	4458
Qy	708	-----LeuProValAspPheValAsnGlyLys	716
Db	4459	TATAATGCTGSCATTCAGTTGATTATACATTTTACTCCGACAGATAATATCAGCGTTAAG	4518
Qy	717	ProSerGlnGluAlaLeuLeuAlaGlyValAla-----ValHis	729
Db	4519	CCTATTTCCTCGCAATATGTTGATGTTTTTCAACCGTACCGTACAAACCCAGCTAAAT	4578
Qy	730	LysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeu	749
Db	4579	ACGCGGTG-----ITGCAACAACCATTTGGACGTTATTGGCAAAA-----	4620
Qy	750	GluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIle	769
Db	4621	GAAGTGGGATTAAGAAGCGGAAATTTTACATTTCCAACTTCTC-----	4662
Qy	770	SerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThr	789
Db	4663	-----GCTTTTATTCTAAATCTCAAGGTTCCCACTCGGCAACACGCAAAAT	4710
Qy	790	GlyGlyIleGlnAlaGlyTyrIleTrp	798

```

Db 4711 GTGGGCTGAATGGGTATCGTTGG 4737
|||||..... ||||| ||
RESULT 4
US-09-815-242-4580
; Sequence 4580, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cart, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4580
; LENGTH: 7104
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4580

Alignment Scores:
Pred. No.: 0.00294 Length: 7104
Score: 137.50 Matches: 204
Percent Similarity: 29.88% Conservative: 142
Best Local Similarity: 17.62% Mismatches: 425
Query Match: 2.91% Indels: 387
DB: 10 Gaps: 47

US-09-914-168-2 (1-919) x US-09-815-242-4580 (1-7104)
Qy 2 SerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeu 21
||||| ||||| ||||| ||||| |||||
Db 1411 ACTAAGAAGTTTATTCCAGATAATCT- - - - -TTA 1443
Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIle- - - - - 38
||||| ||||| ||||| ||||| |||||
Db 1444 AAATTTATCATATAAGTTAATGTTCGGAATATCGATACACCTTAAATATTTGATTTTAA 1503
Qy 39 - - - - -IleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsn 56
||||| ||||| ||||| ||||| |||||
Db 1504 GAAAAATTTACATATCGTACTGCTTCAGATATTGTAATTAATGATCGCCAA- - - - - 1554
Qy 57 ProProValLeuLeuThr- - - - -ProGluGlnIleGlnAlaArgLeuAsnAlaGly 74
||||| ||||| ||||| ||||| |||||
Db 1555 CCAGAAGTTACCTAACTGCAGATCCATTTTTCAGTAGCAGTGAATGAACAAAGATCG 1614
||||| ||||| ||||| ||||| |||||
Qy 75 LeuAsnAlaLysProGlnSerGlnAla- - - - - - - - - - - - - - - - - - - 83
||||| ||||| ||||| ||||| |||||
Db 1615 TTGCACAAACAGTAACTCAAGAGTTGATAATAGTCATTACACAAACAGCATCAATTGCA 1674
||||| ||||| ||||| ||||| |||||

```

Qy 84 -----Leu 84  
Db 1675 GAATACAAATAAACTTAAACAACAGCAGACAATATTTTAAATGAAGATGCGAATCATGTT 1734  
Qy 85 AspValValAsnPheAspAspGlnSerProile  
Db 1735 GAACTGCAATATCGGCATCTCAAGCGCTATTGATGTTAGTAACTAAATTTACAAGCT 1794  
Qy 96 SerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGluGluThr 115  
Db 1795 GCATTAATTGATTAATCAAGCAGCAATTCCTGAATTAGATGCTAAAGCTCAAGAAAGGTT 1854  
Qy 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSer-----ThrGlu 130  
Db 1855 ACAGCAGCACACAAGTAAAAAGTTACGCAAGATGAAGTTGCGACGCTTGTGACTAAA 1914  
Qy 131 MetGlyIleAsnProAsnAspTyrIleProGluTyr-----GlnGly 144  
Db 1915 ATTACAAATGATAAAATAATGCAATCGCAGAAATTAATAACAACAACAGCAGCAAGGT 1974  
Qy 145 -----GluGlnProAsn-----SerGluValValValProProThr 156  
Db 1975 GTTAACTGAAAAGATAACGGTATCGCAGTGTGTAGATCAAGATGTGATTACACCAACA 2034  
Qy 157 LeuGluProGlu---LysProGlyLeuIleLysArgLeuTyrAlaArg----- 171  
Db 2035 GTTAAACCTCAAGCGAAACAAGATATTATCCAAAGCAGTTACAACTCGTAAACAACAATA 2094  
Qy 172 -----LeuPheAsnAspGlyValAsnLys 179  
Db 2095 AAAAGTCAATGCATCATTTACAGATGAAAAGATGAGCAAAATGATAAAATTTGGTAAA 2154  
Qy 180 ValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly----- 194  
Db 2155 ATTGAACAAGGCAATTAAGATATTGATCGACGACAACAACAATGCACAGTAGAAGCC 2214  
Qy 195 ---GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLys 213  
Db 2215 ATTTAAACAAGGCGATCAATGATATTAAATCAA---ACTGCACCTGCTACACAGCTAAA 2271  
Qy 214 ---AlaAlaLeuGlu----- 217  
Db 2272 GCAGCGCTCTTGAAGAAATTTGACCAAGTTGTCACGACAAATTTGATCAAGCAGCTTTA 2331  
Qy 218 -----AspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArg 235  
Db 2332 AATCTGTATACACAACAATGAAGATAGCGGAA-----GCTATTGAAGTATTAAAT 2382  
Qy 236 GlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle-----AspLeuSer 253  
Db 2383 GCAGCTAAAGTTCTGTTGTTAAAGCAATTTGAAGCGACACGACTGCACAGATTTAGAA 2442  
Qy 254 IleIleArgAsn----- 257  
Db 2443 AGAGTTAAACGAGAAATCTCAAAATTTGAAATATTTACTGACTCTACTCTACAAACAAA 2502  
Qy 257 ----- 257  
Db 2503 ATGGATGCTATATGAAGTTAAACAACAGCAGCAACAGCTAGAAAAACTCAAAATGCTACA 2562  
Qy 258 -----SerIleGlyGluValAspValIleIleHisAspLeuGly 270  
Db 2563 GTTTCAAAATGCACTAAATGAAGAAGTAGCAGAGCTAGCAGCAGTATGAGCAGCTCAAA 2622  
Qy 271 GluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys 290  
Db 2623 AAGCAAGTTTACATGACATCCAACTGTTAAATCAAAACAGGAGTGTGCTGATACAAA 2682  
Qy 291 -----AlaPheThrThrValAlaAspGluValProLeuLeu 302  
Db 2683 TCAAAAGTATTAGATAAATCAATCCAAATTCACACACAGCAAAAGTTAAACCT----- 2736  
Qy 303 IleGlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAla 322

Db 2737 GCAGCTGATAGGAGTAGAAAAACGCATATAATACACGTAAACAAGAAATTTCAAATAGC 2796  
Qy 323 SerAlaGlu-----HisGlyTyrPheAspGlyArgTrpLeu 334  
Db 2797 AATGCTTCAACTACAGAGAAAAACAACAGCTCATATACAGAAATTAGATACTAAAAAGCAA 2856  
Qy 335 AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp 354  
Db 2857 GAAGCAAGCAACAATCTTGTGCTGCTCAAAATACAAACAGCTGATTAACAACAGCTAAAGAC 2916  
Qy 355 ThrGly-----ThrGlnTyrArgPheAsp----- 362  
Db 2917 AATGCTATTGCTGCAATTAATCAAGTACAGCGGCAACAACCTAAGAAATTCGGATCTAAA 2976  
Qy 363 ---GluValValPhePheThrIleAspProLysThr-----AsnGlnLeu 376  
Db 2977 GCGGAATTCGCTCAAAAAGCAAGTGAAGTAAACTGCAATTTGAAGCAATGATGATTCG 3036  
Qy 377 ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal 396  
Db 3037 ACTACTGAAGAACAACAA-----GCTCAAAAAGATATAAGTTGATCAAGCAGTAGTTACT 3090  
Qy 397 AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla 416  
Db 3091 GCAACACGCTGATATAGATAAT-----GCTGCGACAAATACTGATGATAGATAAT 3138  
Qy 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu----- 431  
Db 3139 GCAAAACTACTAATGAAGCTACAAATCGCAGCCATTCACCTGATGCGCAATTTTAAACCA 3198  
Qy 432 -----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer 445  
Db 3199 ACAGCGAAACAAGCAATTTGCTGATAAAGTACAAAGCGCAAGAAACAGCAATTCCTAAT 3258  
Qy 446 SerSerSerArgTyrGlu-----ProAlaGlnValAspGluSer 458  
Db 3259 AACGGTGCACAACAAGAAAGAAAGACAGCTCGCAAAACAACAAAGTTTCAAACTGAAAAACA 3318  
Qy 459 ThrLeuGluProValIleGlu-----ThrValGluLeuThrAspGlyIleLeu 474  
Db 3319 ACAGCTGATACAGCAATTTGCTGCATACATAACAATGAGAAAGTTGAAGCGCTAAAAAT 3378  
Qy 475 MetAspIleSerProIleGluPhe-----SerAlaSer 485  
Db 3379 GCAGAAATTTGCTTAAATTTGAAGCAATTCAGCGACCAACAACAACATAATGCGAAA 3438  
Qy 486 AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet 505  
Db 3439 CAAGCAATTTGCTACGAAAGCAATGAACGTAAACACAGCAATCGCTCAACACGCAAGACATT 3498  
Qy 506 ProAspAspArgValIleAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly 525  
Db 3499 ACTGCTGAAGAAATTCAGCGGCAAT-----GCG 3528  
Qy 526 ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn 545  
Db 3529 AATGTAGATAATGCTGTGACACACAGCAAAATAACAACATTTGAAGCTGCTAATAGTCAAAAT 3588  
Qy 546 GluVal-----IleAspLeuProGluArgThrAlaLeuAla 557  
Db 3589 GATGTAGACCAAGCGAAACAACACTGGTGAAGCTAGTATTGATTAAGTAAACACCACTG 3648  
Qy 558 AsnArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyrValPheVal 577  
Db 3649 AATAAAAAGCAACAGCAGTTACAGACCGGAAAAATAATATT-----ACAGCTGCA 3699  
Qy 578 AlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrTrpGlySerAspThrGlyThr 597  
Db 3700 ACTGATGATAATGTTGTTAGATACAGCGAAAGATGAGGTAAAAATTCGATTCAAAAGTACA 3759  
Qy 598 ArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla----- 615

```
Db 3760 CAACGACGACGCGGTAAATCAATCAATGCTAAATAAT---GATGTTGATCAAGCTGTGACA 3816
QY 616 -----GlyAlaGluLeuArgLeuSerGluAspLys 625
Db 3817 ACTCAAAATCAAGCAATGTATATACAACTGGTGCT-----ACAAGTGAAGAAAAA 3867
QY 626 LysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg 645
Db 3868 AATGCACCAAGAGATTAGTTTAAAGCTAAAGAAAAAGCGGTATCAAGATATCTTAAAT 3927
QY 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665
Db 3928 GCA-----CAACCAACTAACGAT-----3945
QY 666 ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArg 685
Db 3945 -----3945
QY 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGlu-----703
Db 3946 -----GTTACGCAAAATTAAGATGTTGCGAAAGATGAATTAGCAACAAAGCA 3984
QY 704 -----ThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro 717
Db 3985 GGTATTACTGCAGATACAAACAATTAAGATGTTGCGAAAGATGAATTAGCAACAAAGCA 4044
QY 718 ---SerGlnGluAlaLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeu 736
Db 4045 AGGGAACAAAGACATTATTGCACAAACTGCGATGCGACTACTGTAAGAAAAAGAACAA 4104
QY 737 ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeu 756
Db 4105 GCAAAATCAACAA-----4116
QY 757 ValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGly 776
Db 4117 ---GTAGATGCAGATTACCGCAGGTAAATCAAAATATTGAAATGACAGCTCAATCGAT 4173
QY 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyr 796
Db 4174 GATGTAAACACTGCAAGAGATAATGCAATTCACGAATTTGACCAATTCAGCATCAACA 4233
QY 797 IleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAsp 816
Db 4234 GATGTTAAACGAAT-----GCAAGACGGAATTTGCTAACTGAAATGCAA 4278
QY 817 GlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeu 836
Db 4279 AATAAAATAACTGAATACTTAAATAATACAGACTACTAATGAAGAAAAGGT-----4332
QY 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp 856
Db 4333 -----AACGAT 4338
QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAsp 876
Db 4339 ATT-----GCACGCTTAGCTGCGATATGAAGAGGTTTAAATAAT-----4380
QY 877 ThrLysIleGlyAlaGlyValGlyValArgTyrTrpAlaSerProValGlyGlnValArgVal 896
Db 4381 -----ATTATGCAACAACTACTACAGCTGATGTAAGTACT 4416
QY 897 AspValAlaThrGlyValLysGlu-----GluGlyAsnProIleLys 910
Db 4417 GCTAAGATACACAGTACAAAAGATTCAACAACCTTCATGCAAACTCTGTTAAG 4470
```

## RESULT 5

```
US-09-815-242-8291
; Sequence 8291, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
```

```
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8291
; LENGTH: 7107
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(7107)
US-09-815-242-8291

Alignment Scores:
Pred. No.: 0.00294 Length: 7107
Score: 137.50 Matches: 204
Percent Similarity: 29.88% Conservative: 142
Best Local Similarity: 17.62% Mismatches: 425
Query Match: 2.91% Indels: 387
DB: 10 Gaps: 47

US-09-914-168-2 (1-919) x US-09-815-242-8291 (1-7107)
QY 2 SerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeu 21
Db 1411 ACTAAGAAGTTTATTCCAGATAAATCT-----TTA 1443
QY 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIle-----38
Db 1444 AAATATCATATAAAGTTAATGTTGCGAATATCGATACACCTAAAAATATTGATTTAAT 1503
QY 39 -----IleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsn 56
Db 1504 GAAAAATTAACATATCGTCTGCTTACATATGTAATTAATAATGCGCAA-----1554
QY 57 ProProValLeuLeuThr-----ProGluGlnIleGlnAlaArgLeuAsnAlaAlaGly 74
Db 1555 CCAGAAGTTACACTAACCTGCAGATCCATTTTCAGTAGCAGTTGAAATGAACAAGATGCG 1614
QY 75 LeuAsnAlaLysProGlnSerGlnAla-----83
Db 1615 TTGCAACAACAAGTAAACTCACAAGTTGATAAATAGTCATTACACACGATCAATTTGCA 1674
QY 84 -----Leu 84
Db 1675 GAATACAATAAATTAACAACAACAGCAGACAATAATTTTAAATGAAGATGCGAATCATGTT 1734
QY 85 AspValValAsnPheAspAspGlnSerProIle-----95
Db 1735 GAAACTGCAAAATCGTCATCTCAAGCGGCTATTGATGTTTAGTAACATAAATACAAAGCT 1794
```



Qy	96	SerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGluThr	115
Db	1795	GCATTAAATGTAATCAAGCAACAATTGCTGAATTAGATGCTAAAGACTCAAGAAAGGTT	1854
Qy	116	ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSer	130
Db	1855	ACAGCAGCACCAAAAGTAAAAAGATTACCGCAAGATGAAGTTGCAGCGCTTGCGACTTAA	1914
Qy	131	MetGlyIleAsnProAsnAspTyrIleProGluTyr	144
Db	1915	ATTAACAATGATAAAAAATATCAATCGCAGAAATTAATAACAAACAAACAGCACAAGT	1974
Qy	145	-----GluGlnProAsn-----SerGluValValProProThr	156
Db	1975	GTTACAACTGAAAAAGATAACGGTATCGCAGTGTGTAGATCAAGATGTGATTACACCAACA	2034
Qy	157	LeuGluProGlu-----LysProGlyLeuIleLysArgLeuTyrAlaArg	171
Db	2035	GTTAAACCTCAAGCGAAACAAGATATTCCCAAGCAGTTTACAACCTGTAANCAACAAT	2094
Qy	172	-----LeuPheAsnAspGlyValAsnLys	179
Db	2095	AAAAAGTCAAATGCATCATTTACAAGATGAAAAAGATGTAGCAAAATGATAAAATTGGTAA	2154
Qy	180	ValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly	194
Db	2155	ATTGAACAAAGGCAATTAAAGATATTGATGCAGCAACAACAANAATGCACAGTAGAAGCC	2214
Qy	195	---GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLys	213
Db	2215	ATTAACAAACAAAGCGATCAATGATATTAAATCAA---ACTGCACCTGCTACACAGACTTAA	2271
Qy	214	---AlaAlaLeuGlu-----	217
Db	2272	GCAGAGCTCTTCGAGAATTTGACGAAGTTGTTCAAGCACAAATTTGATCAAGCACCTTTA	2331
Qy	218	-----AspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArg	235
Db	2332	ATCCCTGATACACAANTGAAGTAGCGGAA-----GCTATTGAACGTATTAT	2382
Qy	236	GluThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle-----AspLeuSer	253
Db	2383	GCAGCTAAAGTTTCTGGTCTTAAAGCAATTTGAAGCGACAACGACTGCACAAGATTTAGAA	2442
Qy	254	IleIleArgAsn-----	257
Db	2443	AGAGTTAAAAACGAGAANAATCTCAAAAAATTGAAAAATATTACTGACTCTACTCAAAACAAA	2502
Qy	257	-----	257
Db	2503	ATGGATGCCTATATCAAGTTTAAACAGCAGCAACAGCTAGAAAAAATCTCAAAATGCTACA	2562
Qy	258	-----SerIleGlyGluValAlaAspValIleIleHisAspLeuGly	270
Db	2563	GTTTCAAATGCAACTAATCAAGAAGTAGCAGAAGCTGATGCAGCAGTAGAAGCAGCTCAA	2622
Qy	271	GluProValTyrIleAspTyrArgAlaValGluValArgGlyGlyGlyAlaAspAspLys	290
Db	2623	AAGCAAGGTTTACATGACATCCCAAGTTGTTAATCAAAACAGGAGGTGTCTGATACAAAA	2682
Qy	291	-----AlaPheThrThrValAlaAspGluValProLeuLeu	302
Db	2683	TCAAAAGATTAGATAAAATCAATGCAATTTCAACACCAAGCAAAAGTTAAACCT	2736
Qy	303	IleGlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAla	322
Db	2737	GCAGCTGATACGGAAGTAGAAAAACGCATATATACACGTCAACCAAGAAATTTCAAAATAGC	2796
Qy	323	SerAlaGlu-----HisGlyTyrPheAspGlyArgTyrLeu	334
Db	2797	ATGCTTCAACTACAGAAGAAAAACAAAGCTGCATATACAGAATTAGATCTACTATAAAAGCAA	2856

Qy	335	AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp	354
Db	2857	GAACCAAGCAAAATCTTGATGCTGCAAAATACAAACAGTGATGTAAACACAGCTAAAGAC	2916
Qy	355	ThrGly-----ThrGlnTyrArgPheAsp-----	362
Db	2917	AATGGTATTGCTGCAATTAAATCAAGTACAAAGCGGCAACAACCTAAGAAAATCGATGCTTAA	2976
Qy	363	---GluValValPhePheThrIleLeuAspProLysThr-----AsnGlnLeu	376
Db	2977	CGCGAAATCGCTCAAAAGCAAGTGAACGTAAAACTGCAATTGAAGCAATCAATCAATTCG	3036
Qy	377	ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal	396
Db	3037	ACTACTGAAGAACAACAA-----GCTGCAAAAGATAAAGTTGATGACAGCAGCTAGTACT	3090
Qy	397	AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla	416
Db	3091	GCAAACGCTGATAGATAAT-----GCTGCAGCAAAATACTGATGTAGATAAT	3138
Qy	417	ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu-----	431
Db	3139	GCAAAACTACTAATGAAGCTACAACTCGCGCCATTACACCTGATGCAAAATGTTAAACCA	3198
Qy	432	-----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer	445
Db	3199	ACAGCGAAACAAGCAATTTGCTGATAAAGTACAAAGCGCAAGAAACAGCAATTCATGCTAA	3258
Qy	446	SerSerSerArgThrGlu-----ProAlaGlnValAspGluSer	458
Db	3259	AACGGTGCACACACAGCAAGAAAAGCAGCTGCGAACAACAACAAAGTTCAACCTGAAAAAACA	3318
Qy	459	ThrLeuGluProValIleGlu-----ThrValGluLeuThrAspGlyIleLeu	474
Db	3319	ACAGCTGATACAGCAATTTGATGTGTGCACATACAAATGCGAAGATTGAAGCGCTAAAAAT	3378
Qy	475	MetAspIleSerProIleGluPhe-----SerAlaSer	485
Db	3379	GCAGAAATGCTATAAATTGAAGCAATTTCAGCCAGCAACAACAACATAATGCGAAA	3438
Qy	486	AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet	505
Db	3439	CAAGCAATTGCTACGAAGCGGAATGAACCTAAACAGCAATGCTCAACCCAGACACATT	3498
Qy	506	ProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly	525
Db	3499	ACTGCTCAAGAAATTCAGCGGCAAT-----CGC	3528
Qy	526	ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn	545
Db	3529	AATGTAGATAATGCTGTGACACAAAGCAAAATAACAACATTTGAAGCTGCTAATGATCAAA	3588
Qy	546	GluVal-----IleAspLeuProGluArgThrAlaLeuAla	557
Db	3589	GATGTAGACCAAGCGAAACAACCTGCTGAAGCTAGTATTGATCAAGTAAACCAACACAGTT	3648
Qy	558	AsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheVal	577
Db	3649	AATAAAAACGACACAGCAGTTCACAGCGCAAAAATAATATT-----ACAGCTGCA	3699
Qy	578	AlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPglySerAspThrGlyThr	597
Db	3700	ACTCATGATAATGGTGTAGATACAGCGAAAGATGCAGGTAAAAATTCGATTCAAAGTACA	3759
Qy	598	ArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla-----	615
Db	3760	CAACCAAGCAACAGCGGTAAATCAAAATGCTAAAAAT---GATGTTGATCAAGCTGTGACA	3816
Qy	616	-----GlyAlaGluLeuArgLeuSerGluAspLys	625
Db	3817	ACTCAAAATCAAGCAATTGATAATCACTGGTGCT-----ACACTGAAGAAAA	3867
Qy	626	LysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg	645



QY 155 oThrLeuGluProGluLysProGluLeuLeuLysArgLeuTyAlaArgLeuPheAsnAs 175  
Db 23120 GCATTATCGCGAGTCACTGCGGGCAGCTGGAGCGCTCAGTCGTCTATTCGGGNA 23179  
QY 175 pGlyValAsnLysValProArgLeuLysAlaLysPheTyArgGlnSerSerGlnSerGlyCl 195  
Db 23180 TGGTATACGCAAA-----CATGCTGCGCGGTGCGCT 23212  
QY 195 uThrSerAlaIleGlySerSerHisGlnLysThrGluPro-----TyAlaAsnI 212  
Db 23213 GCCTCAGTGGTGGGAGTGGTGCATTTCTGGTGAACGATATCGGTATGGTGGCGAGAGC 23272  
QY 212 eLysAlaLeuLeuGluAspIleThrGlnLysSerAlaMetAspLeuAsnGlySerIlePr 232  
Db 23273 TGCCGCACAGATGGAGCGATCGTTGGCCAGTCGTCAGCGATACATAAATCAGTTAA 23332  
QY 232 oArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyTyArgPheAspLe 252  
Db 23333 GCGGCTGAAGGATGATCTGTAAATGCGCGGAAGCT----- 23369  
QY 252 uSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluPr 272  
Db 23370 -----CTGGACAATGAGCTGCATTTCTTACTGCCAC 23401  
QY 272 oValTyIleAspTyArgAlaValGluValArgGlyGluGlyAlaAspAspLys----- 290  
Db 23402 TCAGCTTGAGCAGATACGCTCTTGGGATGCGGCCATGAATATTTGGTGAAGATCGCTGGA 23461  
QY 291 -AlaPheThrThrAlaAspGluValProLeuLeuIleGlyAspValPheHisGln 310  
Db 23462 ACCCATGTCTGACATGGCAGAGAACCGGTCGCGGTACTGCGGATATT----- 23510  
QY 310 yLysTyArgGluThrLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyPheAs 330  
Db 23511 -----GATAATAACCTCAATCGCTTGGCAGTACGCTCAAGTATCTGTC 23554  
QY 330 pGlyArgTrp-----LeuAs 335  
Db 23555 TGATTTGTGGAGTCGTTCTGGGATGCGGCCATGAATATTTGGTGAAGATCGCTGGA 23614  
QY 335 pArgSerValAspValIle-----LeuProAspAsnTh 346  
Db 23615 TGAACAGATTCGCGTTTACAGGAGAAAGTGTGCGGGGGAAGAACTCCCTGGACGCGC 23674  
QY 346 rAlaAspValSerLeuIleTyArgThrGlyThrGlnTyArgPheAspGluValAlaPh 366  
Db 23675 ATCATCTTCTCAGGTTGAATACGAT-----CAGCAGCGCTTTAAGCAT----- 23717  
QY 366 ePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVally 386  
Db 23718 -----CTTCAGGAGAGAAAACGCCAGAGGATTTGCGAGTGCAGAAAGAGCGCAGA 23770  
QY 386 sArgGluLeuLeuGluGln-----LeuLeuThrVa 396  
Db 23771 GCGGAATATACGAGAGCAACAGAAACGCGGTAATGCTGAAATGCTGCACCTAACCGGAT 23830  
QY 396 lAsnMetGlyGluAlaTyArgAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAl 416  
Db 23831 GAATCAACGGAAGCAGCAGCATCAGCGTGAATAGCGCGTATTAAT-----GC 23881  
QY 416 aThrArgTyPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGl 436  
Db 23882 CATGAGTACGCGCATAGGCTGTGAGGATGCGGCG-----ATACA 23923  
QY 436 nAsnAspGlnValSerPheGluGln-----SerSerSerArgThrGluProAl 453  
Db 23924 ACGTGAATAATGACCTTACGAGAAAGCCCTGGCATTCCTGGTAAAGAAAAACAGCGGAAC 23983  
QY 453 aGlnValAspGluSerThr-----LeuGl 461  
Db 23984 CCGTAATGATGAGGCCACCCCGTTATTGCTGCAGTACAGTACAGCAACAGGCACAGGTGA 24043

QY 461 uProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGl 481  
Db 24044 AGGACAGATTGCTGTCGCAGACAGTCAGCAGGCATT----- 24080  
QY 481 uPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHi 501  
Db 24081 ----GCCACGGAAGGATGACAGAACGCGTAAACAGCTTCTGGCTCTGCACAGCGC-- 24134  
QY 501 sLeuTyArgMetProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnAr 521  
Db 24135 -ATCAGCACCTGGAGCGGAAAAAAGTACGGCA-----GATGAA-----AA 24175  
QY 521 gSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAs 541  
Db 24176 GAGTGTGTCGCCCTTAAAGATGAATGATTCAGGCACCTGACG-----CTGCTGGA 24226  
QY 541 pClnSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysTh 561  
Db 24227 TGTAAACAGCAGGAG-----CTTCAGAAACAGACGCGCACTCAACGAGCTGAAG-- 24275  
QY 561 rProAlaAspValTyArgGlnSerLysValProLeuTyArgPheValAlaSerAspLys 581  
Db 24276 -----AAAAACAATTCAGCTGACCTGACCTCACTGGCTGAAGAGA 24316  
QY 581 sProArgAspGlyGln-----IleGlyLeuGlyTyArgLys 593  
Db 24317 GCGCGCTCAGGCTCAGCAACATGACCTGATATGCCACGCTGGGTATG----- 24365  
QY 593 rAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTy 613  
Db 24366 -----GTTGATCAGCAGCGACGATATCAGGTACACTGAGTCTTCGCCAGAAATA 24418  
QY 613 rGlnAlaGlyAlaGlu-----LeuArgLeuSerGluAspLysLysGlyValLysLeuTy-- 631  
Db 24419 CCAGCAACAGCTGGAGCAGTGTGAGCGGATAGTCAGCAGAAAGGACA-----TATAA 24472  
QY 632 -----AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuAr 645  
Db 24473 CACGGATGACTACAGAAAGCGCAGCGCTGACGAGAGCTGACCGCAACATGAA 24532  
QY 645 gAlaThrLeuGlyTyArgGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSe 665  
Db 24533 TCAGAAATCGCGTTTACTGGCAACACGCTT----- 24560  
QY 665 rThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyArgAsnAr 685  
Db 24561 -----CAAGTTGTGTACAG-----GGTAACCTGG--AA 24583  
QY 685 gThrTySerSerLeuArgTyArgLeuAspLysLeuLysThrGlnAlaProProGluThrTr 705  
Db 24584 AAACGAGCTCCTGCT-----GCATT 24604  
QY 705 pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGl 725  
Db 24605 TCAGGATTTTACCGTGGATGTCAGATATACGGCAGAAACAGCAGAA----- 24650  
QY 725 yValAlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyTyArgGl 745  
Db 24651 ----CAGGTGTCTCGTCAGCGCTTCAGCAACATGGAATGCGCTGGCACTTTTCTCAC 24706  
QY 745 nArgTySerLeuGluValGlySer-----SerGlyLeuValSerAspAlaAsnMetAl 763  
Db 24707 TACCGGCAAACTCAATTTCAAACTCTTCACTCTTCTGCTGTGCAGATATGCGGAAAT 24766  
QY 763 aIleAlaArgAla-----GlyIleSerGlyValTySerPheGl 776  
Db 24767 CTGCGCGCAGGCAACCATGATGAATCGATAAAGGATTTGCGAGTGTACTGGGATTGA 24826  
QY 776 yAspAsnAlaTyArgGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTy 796  
Db 24827 TCTCAGACGCTTTCCTCTGAAT-----GCCAATGGGGGATTATATCAGTCTGC 24874  
QY 796 rIleTrpSerAspAsnPheAsnHisValProTyArgLeuArgPhePheAla----- 813

```
Db 24875 TGATTGAGTCGTTACAGTGGCACCGTGGTTAAACGCTCCGAGCTTTTGTGTTTGTGAAA 24934
Qy 814 -----GlyGlyAspGlnSerIle-----ArGlyTyrAl 823
Db 24935 AGCGCGGGTGTGATGGGGAAGCGGACCTGGAACCAATCTGCATTCGCTGCGTGGTGC 24994
Qy 823 aHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAl 843
Db 24995 TGACGCTAAGCTGGGGTGTGGCGAT-----ATTGGGGTTCAGGTATGGC 25042
Qy 843 aValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGl 863
Db 25043 GATGTTTCCCGCGAGTAAACATCGAGATCAATAACGAT-----GCCACGAACGG 25093
Qy 863 yAspIleGlyAsnAla-----TyrAspLysGly 872
Db 25094 GCAGATAGTCCGGCTGCCTGAAGCGGTTTATGACCTCGGG 25136

RESULT 7
US-09-815-242-7474
; Sequence 7474, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7474
; LENGTH: 8673
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(8673)
; US-09-815-242-7474

Alignment Scores:
Pred. No.: 0.0057 Length: 8673
Score: 136.00 Matches: 198
Percent Similarity: 33.56% Conservative: 145
Best Local Similarity: 19.37% Mismatches: 361
Query Match: 2.88% Indels: 318
DB: 10 Gaps: 52

US-09-914-168-2 (1-919) x US-09-815-242-7474 (1-8673)

Qy 55 GlyAsnProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaGly 74
```

```
Db 5023 GGAGCGCCAGAAATCATTTGCGCAATGAAAA----- 5055
Qy 75 LeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGlnSerPro 94
Db 5056 -----AGGATGTTGCAAGAGCGGTGATGCTGCTTTTATACACGCCGTAGCAC 5106
Qy 95 IleSerArgIleGlyGluGlnSerProLeuGlyLeuAspMetSerValIleGluGlu 114
Db 5107 AATGCGGTTAAAGGGGTAACAAACGCCCTTTA----- 5139
Qy 115 ThrThrProLeuSerLeuGluGluLeuPheAlaGlnGlnSerThrGluMetGlyIleAsn 134
Db 5140 -----AAATCGCTCAGCGAGATCATTAAGGCAAGCAAGGCGCTTTCAGCGCAAAAC 5190
Qy 135 ProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValValPro 154
Db 5191 CTTTATAGTAAGCGCGTGGATTTTTCAGGC-----AGAAGCGTGTATGTCGTGGG 5241
Qy 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174
Db 5242 CCTAACCTTAAATGATGATCGGATTCCTTAAAAACATGCGTGAACACTTTCAAA 5301
Qy 175 AspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly 194
Db 5302 -----CCGCATCTGTTATCCACGCTT-----GAAGAGACAGCG 5334
Qy 195 GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsn-----IleLys 213
Db 5335 TATGCCACCACGCTCAACAGCGCTAAACAGCTGATTTGACAAAGAGCAATGAAGTGTGG 5394
Qy 214 AlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArg 233
Db 5395 GAGTGTCTGCAAGAAATCAGAGGGGTATCCGGTGTCTACTCAACCGGCC-----CCTACC 5451
Qy 234 LeuArgGlnThrAla-----LeuValAlaAlaArgAlaValGly 246
Db 5452 TTGCACAAGCAATCCATCAAGCGTTCCATCCAAAGCTCATTCAGCGCAAGCGATCCAG 5511
Qy 247 TyrTyr-----AspIleAsp----- 251
Db 5512 TTGCACCCGTTAGTGTGCTCAGCGTTTAAAGCGTGTATTTGACGGGGATCAATGCGGTG 5571
Qy 252 ---LeuSerIleIleArgAsnSerIleGlyGlu-----Val 262
Db 5572 CATGTGCTTTAAGCCAGGAAGCGATCGCTGAATGCAAGGTCTCATGCTAAGCTCTATG 5631
Qy 263 AspValIleIleHisAspLeuGlyGluProValTyrIleAsp----- 276
Db 5632 AATATCCTTTTACCCTAGCGGTAAAGCGGTAGCCATTCCTTAGCCAGGATATGCTTTTA 5691
Qy 277 -----TyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThr 293
Db 5692 GGGCTTTATTATCTTTTACAAAAGAGTGGGTCAAGGGCGAGCATACGCTTTCCT 5751
Qy 294 ThrValAlaAspGluValProLeuLeu-----IleGlyAspValPheHisGlyLysTyr 312
Db 5752 AGCGTGAATGAATCATCACCGCTATTGACACGAAGAAGATTAGACATCCACCAAGATT 5811
Qy 313 GluThr-----LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAsp 330
Db 5812 AGGGTTTGGATCAAGGGAATATTATTCGCTACGACGCGG----- 5850
Qy 331 GlyArgTyrLeuAspArgSerValAspValIleLeuProAsp----- 344
Db 5851 GGGCGCATGATCATCAAGTCC-----ATTTCCTGCTGATTTATCCCTACCGATTG 5901
Qy 345 -----AsnThrAlaAspValSerLeuIleTyrAsp----- 354
Db 5902 TGAACACACCCATGAAGAAAAAAGATATTGGCGTGTGTTGGATTATGTCATAAAGTG 5961
Qy 354 ----- 354
```

Db 5962 GCGGTATCGCATACCGCAACCTTTTGGATCANTTTGAAACGCTTGGCTTTAGGTAT 6021  
Qy 355 -----ThrGlyThrGlnTyArgPheAspGluValValPhePheThrIleAspPro 371  
Db 6022 CGCACTAAGCGTGGTATTTCTATCTCTATGGAAGATATT-----ATCAGCGCA 6069  
Qy 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArg-----GluLeu 389  
Db 6070 AAAGACAAGCAAAATAATGGTGGAAAAAGCCAAAGTAGAAGTTAAAAAATCCAACACAA 6129  
Qy 390 LeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyAsnLeuGlnAlaValArgAla 409  
Db 6130 TACATCAGGGTGTCTCTACCTGACCA--GAGCGTTAC----- 6165  
Qy 410 LeuSerAsnAspLeuIleAlaThrArgTyPheAsnMetValAsnThrGluIleValPhe 429  
Db 6166 -----AATAAAATCATTGACACT-----TGGACTGAAGTCAATGACAGAATGAGTAAG 6213  
Qy 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln----- 444  
Db 6214 GAAATGATGAGCGGTATCGCAAAAGATAAAGAGGCGTTTAAACTCTATTATATGATGCG 6273  
Qy 445 SerSerSerArgThrGluProAlaGlnVal----- 455  
Db 6274 GATAGTGGCGCAGGGGAGCGCGCGCAATCCGTCAGCTTTCAGCGATGAGGGGGCTT 6333  
Qy 456 -----AspGluSerThrLeuGlu---ProValIleGluThrValGluLeuThr 470  
Db 6334 ATGACAAAGCGGATGGCAGTATCATTTGAAAGCGCCATATTCTTAACATTTAAAGAGGG 6393  
Qy 471 AspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAsp 490  
Db 6394 TTGAATGCTTTAGAAATCTTTAACTCCACGATCGCGTAGAAGGGCTTAGCGGATACA 6453  
Qy 491 LysLeuAsnLeuValAlaAla-----LysAlaArgHisLeuTyArgMetProAspAsp 508  
Db 6454 CGCTAAACAGCAATGCGAGGTATTTTGACAGAAACATCATTCAGCTT---TCGCAA 6510  
Qy 509 ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528  
Db 6511 AATGCTCAAGGTGGTCTCTGATGATTGCGGCACGCGATGAAGGATT-----GAAATCACG 6564  
Qy 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu----- 544  
Db 6565 GAT---ATTGCGGTGGGAGTGAGCTGATGAACCTTTAGAAGAGCGTATTTTCGGGCGC 6621  
Qy 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561  
Db 6622 GTTTTATTAGAGATGTGATCGATCCCATTTACCAATGGAATCTTGCTCTATCGGACACT 6681  
Qy 562 ProAlaAspValTyArgLysLysVal-----ProLeu 573  
Db 6682 TTGATTGACGAAGAGGTGCTTAAAGGTGGTTGAAGCTGGGATTAATCCATTACGATC 6741  
Qy 574 TyrValPheValAlaSerAspLysProArgAsp-----GlyGlnIleGly 588  
Db 6742 CGCACCCAGTAAGTGTAAAGCGCCAAAGGCGGTGTCGCCGAATGCTATGGCTTGAAT 6801  
Qy 589 LeuGlyTyArgGly----- 592  
Db 6802 TTAGCGGAAGCAAGATGAGTATCCGGGTGAAGCGGTGGCGGTAGCCGCGCAATCT 6861  
Qy 593 ---SerAspThrGlyThrArgLeuValThrLysPheGluHis----- 605  
Db 6862 ATCGGGAGCGCTGGAAACCGACTCATCTTTAAGGACTTTTCCATGTGGCGGACACGAGC 6921  
Qy 606 -----AsnLeuIleAsnArgAspGlyTyArgLysGlnAlaGlyAlaGlu 618  
Db 6922 AGGAGTCAGGATGATGCGGAAATCGTGGCGAGCAAGAAGGTTTTGTGCGTTCTTACAAC 6981  
Qy 619 LeuArgLeuSerGluAspLysLysGlyValLysLysLeuTyAlaThrLys----- 634  
Db 6982 CTTAGGACTTACAGAAATAAAGGGTAAACATTTATCGCCAACCCGCGTAACGCTTCT 7041

Qy 635 -----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeu 648  
Db 7042 ATTTTACTGCTAGACCTTAAGATTAAAGCGCTTTTGATGGGAATTACGCATTAAACAG 7101  
Qy 649 GlyTyArgGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThr 668  
Db 7102 GTCTATGAAGAAGTCGTTGTGAGC----- 7125  
Qy 669 LeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyPsnAsnArgThrTySer 688  
Db 7126 -----GTGAAAAATGGCATCAAGACCTTAATTCGTT 7158  
Qy 689 LeuArgTyArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGlnAspLeu 708  
Db 7159 TTAAGA---AGAAGCATATTGTCAG-----CCAAGCGAA----- 7191  
Qy 709 ProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu-----AlaGly 725  
Db 7192 ---TTAGCCGCGCTTGGCGGTAAGATTGAGGGGAAAGTGTATTTCCTTATGCTAGCGG 7248  
Qy 726 ValAlaValHisLys-----ThrValAlaAspAsnLeuValAsnProMetArgGlyTy 743  
Db 7249 CATAGGTGCATAGGGGGGAGTATCGCTGATATTATC-----CAACAGGCGTGG 7299  
Qy 744 ArgGlnArgTySerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763  
Db 7300 AATGTGCCTAATCGCATCCCTTATCGAGCGAATTCGTGTTAAAGAT---AATGACCC 7356  
Qy 764 IleAlaArg-----AlaGlyIleSerGlyValTySerPheGlyAspAsnAlaTy 780  
Db 7357 ATCGCGCAAGATGTGTATGCCAAAGAAAAGGTGTGATCAAAATAC-----TAT 7404  
Qy 781 GlySerAsnArgAlaHis---GlnMetThrGlyGlyIleGlnAlaGlyTyTrpSer 799  
Db 7405 GTTTTAGAGGCCAATCATTTAGAGCGCACCATGAGTGCAGTCAAAAAGGCGCATATTGTGAGT 7464  
Qy 800 AspAsnPheAsnHisValProTyArgLeuArgPhePheAla-----Gly 814  
Db 7465 GAAAAA-----GGCTTTGTCAGTGGTAGCTGATGATAAT 7500  
Qy 815 GlyAspGlnSerIleArgGlyTy---AlaHisAspSerLeuSerProIleSerAspLys 833  
Db 7501 GGTAGGAAGCGCGTCGCCATTATATCGCTAGGGGTTCTGAGATCTTGATTGATGATAAT 7560  
Qy 834 GlyTyLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlu----- 848  
Db 7561 AGTGAAGTGAGCGCTAATAGTGTATTTCTAAACCCAGCACTTAACACTTTCAAAACGATT 7620  
Qy 849 -----TyAsnTyArgGluPheMetLysAspLeuArgLeuAlaValPheGly 863  
Db 7621 GCCACATGGGATCCCTTACAACACCCCTATCATTTGCGGACTTT----- 7662  
Qy 864 AspIleGlyAsnAlaTyArgAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883  
Db 7663 -----AAAGGTAAGGTGAATTTTGTGGATGTTATCGCAGGGGTT 7701  
Qy 884 GlyVal 885  
Db 7702 ACGGTC 7707

## RESULT 8

US-10-108-605-236

; Sequence 236, Application US/10108605

; Patent No. US20020160934A1

; GENERAL INFORMATION:

; APPLICANT: Broadus, Julie

; APPLICANT: Stam, Lynn

; APPLICANT: Bachmann, Jane

; APPLICANT: Kamdar, Kim

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOCASTER THAT ENCOD

; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF

; FILE REFERENCE: 31133B

; CURRENT APPLICATION NUMBER: US/10/108,605  
 ; CURRENT FILING DATE: 2002-03-27  
 ; PRIOR APPLICATION NUMBER: US 07/61,142  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/176,418  
 ; PRIOR FILING DATE: 2000-01-14  
 ; NUMBER OF SEQ ID NOS: 361  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 236  
 ; LENGTH: 12404  
 ; TYPE: DNA  
 ; ORGANISM: Drosophila melanogaster  
 US-10-108-605-236

## Alignment Scores:

Pred. No.: 0.0094 Length: 12404  
 Score: 136.00 Matches: 132  
 Percent Similarity: 31.54% Conservative: 103  
 Best Local Similarity: 17.72% Mismatches: 253  
 Query Match: 2.88% Indels: 257  
 DB: 9 Gaps: 32

US-09-914-168-2 (1-919) x US-10-108-605-236 (1-12404)

Qy 14 ProValAlaLeuAlaAlaValLeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsn 33  
 Db 1966 CCGACTATAATGCTGCTCTCCTGAGTGCAGTCCCTCCAAAATATCGAGTCGGCC 2025  
 Qy 34 AsnProAlaAsnIleAlaAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLys 53  
 Db 2026 CAACCGGGGGAATCGTGGCCAGAATTCGGTGCATGAC----- 2064  
 Qy 54 AlaGlyAsnProValLeuProGluGlnIleGlnAlaAaArgLeuAsnAlaAla 73  
 Db 2065 -----CCCGATTGCAAAACGGAGTATGCCAATGTGAATGTGACGCTCAATGAGGA 2115  
 Qy 74 -----GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPhe 89  
 Db 2116 GATGTCCTCTCGCGTGTGACCAACGAGGATAACAGTATTATTGGTTATAGTTCACCTG 2175  
 Qy 90 Asp---AspGlnSerProIleSerArgIleGlyGlnSerProProLeuGlyLeuAsp 108  
 Db 2176 CCGCTGACAGGGAGATGTTTCA-----AATTATACA 2208  
 Qy 109 MetSerValIle-----GluGluThrProLeuSerLeuGluGluLeuPheAla 125  
 Db 2209 CTGAGTGTGTAGCCACCGATAAGGAACTCCGCCCTGACGCCCTCCAATCGATCTC 2268  
 Qy 126 GlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGlu 145  
 Db 2269 CTAAGGATTACCGATGTCATGCAATCCA-----CCGGAGTTCGAGCAGGAC 2316  
 Qy 146 GlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeu--- 164  
 Db 2317 CTCTACACCGCCACGTC-----ATGAGGTGGCCGATCCGGGTACATCG 2361  
 Qy 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsn----- 178  
 Db 2362 GTACTTCAAGTTCTAGCCACCGATGAGATGAGGCGCTAAACTCGCTTTGACATACTCG 2421  
 Qy 179 -----LysValProArgLeuLysAlaLysPheTyrGln---SerSerGlnSerGlyGlu 195  
 Db 2422 CTGGCCGAACCGCCGAACCCATGCCAGTGGTTTCAGATTGACCCACAAACGGCGCTG 2481  
 Qy 196 ThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAla 215  
 Db 2482 ATCACCACCGGATCGCATATCGATTGCGAAACGAGCGCGTCCGCGAGCTAACGGTTGTC 2541  
 Qy 216 LeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArg 235  
 Db 2542 GCTAGGGAC-----GGAGGTGTCGCCGCCCTTTCC 2571  
 Qy 236 GlnThrAlaLeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleIle 255

Db 2572 TCCACAGCCACGGTCTT----- 2589  
 Qy 256 ArgAsnSerIleGlyGluValAspValIleIleHisAspLeu-----GlyGluProVal 273  
 Db 2590 -----GTTACCATTCCAGTGTAAACGACACGACACCGCAT 2625  
 Qy 274 TyrIleAspTyrArgAlaValArgGluGlyAlaAspAspLysAlaPhe--- 292  
 Db 2626 TTC-----GACCAGAGCTTCTAC 2643  
 Qy 293 ---ThrThrValAlaAspGluValProLeu----- 301  
 Db 2644 AACGTATCCGCTCGCTGAAACGACCGTGGCAGGTGCATCTCAAGTCTCTGCAAGT 2703  
 Qy 302 -----LeuIleGlyAspValPheHisHis 309  
 Db 2704 GATCCAGATTGGGAGTAATGCCATGTGAACCTACACCATCGCGGAGGCTTCAAGCAC 2763  
 Qy 310 ---GlyLysTyrGluThrLysAsnLeuIleGluAsnAlaSerAlaGlu----- 325  
 Db 2764 CTCACCGAATTTCGAG-----GTGGCTCCGCTTCTGCGAGATTTCGATA 2808  
 Qy 326 HisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsn 345  
 Db 2809 GCTGGCGAATTTCGAC----- 2823  
 Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365  
 Db 2824 -----TTCGAGCGGAGGAGTAGCTATGAATTTCC----- 2853  
 Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385  
 Db 2854 -----GTATGGCAACCGATCGTGGGGTCTTGAGCACT 2886  
 Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetClyGluAlaTyrAsnLeuGln 405  
 Db 2887 ACTGCCATGATTAATGCAATTGACGGAGCTGAAC----- 2922  
 Qy 406 AlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThr 425  
 Db 2923 -----GACATCGT 2931  
 Qy 426 GluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer 445  
 Db 2932 CCGGTTTCTATCCCGAG-----CAATATAA 2958  
 Qy 446 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465  
 Db 2959 GTATCGCTCAGGAATCCCGAAGGCTCTCACAGCTTCGAGTACTCGGATTGCGCT 3018  
 Qy 466 ThrVal-----GluLeuThrAsp 471  
 Db 3019 GTGGTGGCCACAGATCCGGATTACGGAAACTTTGGCCAGGTCTCATATAGGATTGTGGCC 3078  
 Qy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491  
 Db 3079 GGTACGACGGCGGCATCTCCGGATTGATCGTCCACTGGTCAATATTCGTGTCGCT 3138  
 Qy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeu 511  
 Db 3139 CCGGATATGCTATCCGTTCCGTACCCCAACCCATGAC-----ATGCTA 3180  
 Qy 512 AlaIleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAspAlaVal 531  
 Db 3181 AATATATCGCGCACCGATGCGCGTAAT-----CTAAGGAGCAATGCCACCGCGTG 3231  
 Qy 532 SerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuPro 551  
 Db 3232 GTCTTCTTCAGC-----ATCATCCACGCCATG 3258  
 Qy 552 GluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysVal 571

Db 3259 CAGCGACCGCGGATC-----TTCAAAGGCC 3285  
Qy 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGly-----Leu 589  
Db 3286 AGATACAATTACTGTGCAAGGACACATTCGAGGGCCACGGTTGTGGTTCTGTGATA 3345  
Qy 590 GlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsn 609  
Db 3346 GCTCGGAGTGAGATGTCGCCCATCGAAGTCGGTTAGGTATTCGATATACATCGGGCGAT 3405  
Qy 610 ArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLys 629  
Db 3406 CCCGATGGCTAT-----TTACGATAGAAACCACTCTGGAAC 3444  
Qy 630 LeuTyrAlaThrLysProLysProLysHisProLeuAsnAspGln-----LeuArg 645  
Db 3445 ATACGTATAGTAAAGCCCTGGATCAAGCGAAGTCCCGAGTTCTACTCAACATCCAA 3504  
Qy 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665  
Db 3505 GCCACCTTGGGA---GAACCTCCGGTTATGGACACACAGGTTAATATCGAAGTGGAG 3561  
Qy 666 -ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnAr 685  
Db 3562 GATGTCACGATATGCACCTGAGTTCGAGGCCAGTATGGTTAGGATTTCCGTACCGGAA 3621  
Qy 685 gThrTyrSerLeu 689  
Db 3622 AGTCAGAGCTTG 3634

## RESULT 9

US-10-114-170-243/c

; Sequence 243, Application US/10114170

; Publication No. US20030023075A1

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Valerie

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

; TITLE OF INVENTION: No. US20030023075A1el sequences of E. coli O157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles &amp; Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/114,170

; FILING DATE: 01-Apr-2002

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/453,702

; FILING DATE: 03-DEC-1999

; APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95017

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 243:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5629

; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 243:  
US-10-114-170-243

## Alignment Scores:

Pred. No.: 0.00373 Length: 5629  
Score: 135.00 Matches: 205  
Percent Similarity: 31.34% Conservative: 157  
Best Local Similarity: 17.75% Mismatches: 365  
Query Match: 2.86% Indels: 428  
DB: 9 Gaps: 51

US-09-914-168-2 (1-919) x US-10-114-170-243 (1-5629)

Qy 54 AlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaAArgLeuAsnAlaAla 73  
Db 3289 GCCCGCAATTATCCGGCTATAAGCAAGGGAGCTGCACTAAGGGCACACTTAGCCTTCAC 3230  
Qy 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGlnSer 93  
Db 3229 AACTGGAATGCTCCACTCCAATCCCATATTTATACATTGAGGCAACCAACCAATAGGCT 3170  
Qy 94 ProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGlu 113  
Db 3169 CGCGTTGCCACATTATCAGCGACAAATAATGACGCTTTACGCCGATAAAAACACATTTAAT 3110  
Qy 114 GluThrThrProLeuSerLeuGluGluPheAlaGlnGluSerThrGluMetGlyIle 133  
Db 3109 ACCCTCACG-----ATCAACGTCACCTGATGAGAGTGAT 3077  
Qy 134 AsnPro---AsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValVal 152  
Db 3076 AATCCCTCGACAATCATCAGGTCACCTTTAAGAATGAAAGAGGAGCGGAGTTGTC 3017  
Qy 153 ValProPro----- 155  
Db 3016 GAACCGCGCAGCAAAATACGGATGCATATGGTGTGTCACAAATAAACATGTAAGTCAG 2957  
Qy 156 -----ThrLeuGluProGluLysPro---GlyLeuIleLysArgLeuTyr 169  
Db 2956 GTTCGGGAAGAAATACGATTAGCCGCACGCTGCCAATGTTTTCACACGGATAATT 2897  
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGln 189  
Db 2896 GCGAAATTGCTTAGCGAT-----TCGAGTAGCGCA-----AAATTCAMCAA 2855  
Qy 190 SerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr----- 206  
Db 2854 CTGTTTCCCGATCCAGATACCATTTATTGCTGGCAACAGCAGGCGAGTACTCTGACGCC 2795  
Qy 207 -----GluProTyrAlaAsnIleLys----- 213  
Db 2794 ATCATCACAGACTTTTCATACACACCCGTTAAAGATGATGAAGTGAATTTTGTGCGACCT 2735  
Qy 214 -----AlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIle 231  
Db 2734 GGTGGCTCGCAACTGGACAACACGACC-----GCCACAACAGACAGCTCCGGTATTGTG 2681  
Qy 232 ProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAsp 251  
Db 2680 -----CGGGTGCACCTGACCATTTCAAAAGCT---GGTAGCTATTTCGCTCGAT 2636  
Qy 252 LeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGlu 271  
Db 2635 GCCTCGCTT-----GAGGTGGATAAAATATTCACCATGTCGCTCAGC 2594  
Qy 272 ProValTyrIleAspTyrArgAlaValGluVal-----ArgGlyGluGlyAla 287  
Db 2593 ATCACCCTGGTCCCAACAGGGAACAACTCGGTAAATGACCTTGAATGCGGGTCCGGCAGT 2534

QY 287 ----- 287  
Db 2533 GCGATCGCTAAACAATACAAATATCGTTACCTGACTGCCAGTGTGAAGATGTTTATGGA 2474  
QY 288 ----- 296  
Db 2473 CACCCGTTGCCGGATGAGGATGTGAATTTACCTTGCCAGCCTCCATGACCGGGAACITC 2414  
QY 297 ----- 303  
Db 2413 ACGCTAAGTAGTGAACCGCCGACCGATGCAAAAGGTGATGCCGTGGTCACATTGCGA 2354  
QY 304 GlyAspValPheHisHisGlyLysThrGluThrLysLysAsnLeuIleGluAsnAlaSer 323  
Db 2353 GGCACA-----AAAGCGGGTGAGTTTACAGTTACGGCGAGCCTGACCCAGAAATAATACC 2300  
QY 324 AlaGluHisGlyTyr-----PheAspGly----- 334  
Db 2299 GTTGCTTATCAGCAAGTCACTTTTATTGGGGATACAACAGTGGCGAGCTCCAGCCCGTG 2240  
QY 335 AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu----- 351  
Db 2239 ACTGCCTCATTAATTCATTGTTGCGGGTAACAGTACGGGGAGTACCTTGACGGCAACG 2180  
QY 352 IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePhe----- 367  
Db 2179 ATCCTGGACGGTTACCAAAATCCCTTAAGACCGAGTTGGTCACTTTCCACAGTAACGAT 2120  
QY 367 ----- 367  
Db 2119 GTCACCTCTAAGCGAAACAGAAAGTCAACCACCAATACGCTGGTGCAGGCGACGTAACAATG 2060  
QY 368 ----- 370  
Db 2059 ACCAGCAATATTGCCGGACAAACATAAGCTGCTGGTGCAGCGGAAAGCGCAAGCTTCCGAT 2000  
QY 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys-----ArgGluLeu 389  
Db 1999 AATAAAACGTTTAGTTTATCAGTCTACCGGATCAAAAGTTCGGCGAAGTGAATAAGTATA 1940  
QY 390 LeuGluGlnLeuThrValAsnMetGlyGluAlaTyrAsnLeuGln-----AlaValArg 408  
Db 1939 ACCGGACCGCAAAACGATACCGGTGGCGGAAACATCACGCTACGATACTCGTCCAG 1880  
QY 409 AlaLeuSerAsnAspLeuIleAlaThrArg----- 418  
Db 1879 GACCGGTTTACAATGTAATCGCGGGTCAACGGCTCAGATTAAAGTGGCGACCAACAAC 1820  
QY 419 ----- 422  
Db 1819 AACATTACGATAGCGGATACGGCTTACACCGATAAATAACGGTTATGCGTACGTTAACCTT 1760  
QY 423 ValAsnThrGluIleValPhePrdGluArgGluGlnIleGlnAsnAspGlnValSerPhe 442  
Db 1759 CTGACACCAA-----CCTGGGGTTTATCAGGTG-----ACGGCAACGGCTG 1718  
QY 443 GluGlnSerSerSerArgThrGlu----- 451  
Db 1717 GACAAATAACAGTAGTAGTAGTTGACGTGAATGTGGCAATGGCAAACTCGAGTTAACA 1658  
QY 451 ----- 451  
Db 1657 TCATCGAAACCCAGAAACTACGGTCCATAATAGTAGGGTATTACGCTGACCGCAACGGCG 1598  
QY 452 ----- 463  
Db 1597 AGAAATCGCGGGGTGAATGTATGTCAGCGCAAAATTTATCACCTTTTAGCGTAACGCGCTGAA 1538  
QY 464 IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSer 483  
Db 1537 GGTGCAACGCTAGCAATACAGGGGAAGTCTTACTGACCACTCA-----GGTCAGGCC 1484  
QY 484 AlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyr 503

Db 1483 AAAGTGACGTGACCACTGACAAAGTGAATGCTCTATACCGTTACGGCCATAATGGGCAAA 1424  
QY 504 AspMetPro-----AspAspArgValLeuAlaIleAsnHisAsp----- 516  
Db 1423 GATGTTCCCGTTTCAGAGCCAGGTAAACGGTTTAAAGCAGCATTAACACGGCACAT 1364  
QY 517 ----- 522  
Db 1363 GTTGTGACCGTCGTGGTCTTCCTGACACCATCACCGCCAGCGCATCGATACGACGACC 1304  
QY 523 IleLeuGlyArgIleSerAsp-----AlaValSerAlaValAlaArgAlaIle 538  
Db 1303 ATCATCTTCACGAGTAGAAGATGATTACGGATTCCCGGTTCAAGGTGTCCGATATTAGTCAT 1244  
QY 539 LeuProAspGluSerGluAsnGluValIleAspLeuPro----- 551  
Db 1243 GGCTTAGACACCAAGGACGCGCGTAGTTAATATTCCAACCTACGGGTACCGATCAGTCC 1184  
QY 552 ----- 560  
Db 1183 GGCAAGTCACGGGACCAATAACCAAGTACATTCGGCAGAAACCTTAACAGTCAATGTGCAA 1124  
QY 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580  
Db 1123 GTTCTGGCACAGCCCAACCAATCCGCAACCATTTACATTGTTGCCGCGACGCCGATGAA 1064  
QY 581 LysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuVal 600  
Db 1063 ----- 1052  
QY 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620  
Db 1051 TTGAAATTCGATCTTCACACTCTGAAAGCTGAC-----TACCAGCAGAGCGCAAACTTACG 995  
QY 621 LeuSer---GluAspLysLysGlyValLysLeuTyrAlaThrLysProLeu----- 536  
Db 994 CTAACATTGCAAGACAAGTACGGTAACCCGATAGTAGCTCTGATCATCTGGAATTTGTC 935  
QY 637 ---SerHisProLeuAsnAspGlnLeuArg---AlaThrLeuGlyTyrGlnGluVal 654  
Db 934 CAGTCAGGCCCTTCGTGAACTTCTCAAGTTCAGCGGATATTGATTACAGCCCAAGAAAT 875  
QY 655 PheGlyHisSerThr----- 659  
Db 874 TATGGCGAGTACACCGCTGACTCTACTGGCGGAAAGAGGAAACAGCAGACTCATTTCC 815  
QY 660 ----- 677  
Db 814 ATGCTCAACGGGGTTCATCAGGCAAACTTAAGCATATATCGCTGAATCTCATCAATCGATA 755  
QY 678 IleGlnAsnGlyGlyTrp-----AsnArgThrTyrSerLeuArgTyrArgLeu 693  
Db 754 AAAGAAATGTCGGGTTCATGCTACTGCMAAACACCATACCTCTCTCCACGGCT----- 704  
QY 694 AspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheVal 713  
Db 703 ---AAATCCCGAGCGAAGGCTTTGCAGGAGCGTATTACACACTC----- 662  
QY 714 AsnGlyLysProSerGlnGluAlaLeuLeuAlaValAlaValHisLysThrValAla 733  
Db 661 ----- 626  
QY 734 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753  
Db 625 GATTATATGTTTCAAGTTTCACAGGGTTGG-----GTGCTGCTCCAGTCT 581  
QY 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773  
Db 580 TCGGGTAAAGTTTCTTCGCAAAATATC----- 554  
QY 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791



```
Db 553 -----GGCGATCAACGTCAGTCACAATAAGCGCTGTCCCGACAGAGGAGGTACAACC 500
Qy 792 -----IleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyr 807
Db 499 TACCAGACCTTAATTAAGCTCAAGGCTGGTGGTGAATAATGGAATCATACCAATATC 440
Qy 808 ArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu 827
Db 439 TGGCTA -----GCTGCCAATGCGGCTC 419
Qy 828 SerProIleSerAspLysGlyTyr -----LeuThrGlyGlyGln 840
Db 418 TGTCATGCTAAATAATGATGGATATAATCTTCCTGGCATCACACATTTGACGCTGGGGAA 359
Qy 841 ValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860
Db 358 -----ACAAACGCGACCGGGATCA 338
Qy 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880
Db 337 CTGTATGCTGAATGGGGAACGTT -----GGAGCGTTTTCAGTAATTCGAATTTACA 284
Qy 881 AlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArgValAspValAlaThr 900
Db 283 CCG -----GGAGCTTACTGACAGT ----- 263
Qy 901 GlyValLysGluGlyAsnProIleLysLeuHisPhePheIle 915
Db 262 -----GAATCTGATGATTACAGTCGGCAGCTACTATGTG 230
```

## RESULT 10

```
US-09-815-242-4760
; Sequence 4760, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4760
; LENGTH: 6228
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4760
```

```
Alignment Scores:
Pred. No.: 0.00432 Length: 6228
Score: 135.00 Matches: 154
```

```
Percent Similarity: 32.63% Conservative: 110
Best Local Similarity: 19.04% Mismatches: 301
Query Match: 2.86% Indels: 244
DB: 10 Gaps: 35

US-09-914-168-2 (1-919) x US-09-815-242-4760 (1-6228)
Qy 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAlaAsnIleIleAsnHisValPro 43
Db 4144 GTTCTTCACAAGCATTAGCTAAAGAAAAAGAAAGACACTT----- 4185
Qy 44 AlaHisaspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrPro 63
Db 4186 -----GCAGCTATTGACCAAGCT----- 4203
Qy 64 GluGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsn----- 76
Db 4204 ---CAAACGAATAGTCAGGTGAATCAAGCAGCAACAATGTGTATCAGCGATTAAATTT 4260
Qy 77 -----AlaLysProGlnSerGlnAlaLeuaspValValAsnPheAspGln 92
Db 4261 ATTC AACCTGAAACAAAGTTAAACACAGCTGCAGCTGAAATAATCAAT--- 4308
Qy 93 SerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIle 112
Db 4309 -----CAAAAGCGCAATGAATTAAGTCTAAGATTAAATCAAGATAAAGAACCAACA 4359
Qy 113 GluGluThrThrProLeuSerLeuGlu-----GluLeuPheAlaGlnGluSerThr 129
Db 4360 GCAGAAGAACAGTAGTACACTAGATAAAATCAATGAATTTGTAAATCAAGCCATGACA 4419
Qy 130 GluMetGlyIleAsn---ProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn 148
Db 4420 GATATTACGAATAATAGAACAAATCAACAAGTTGATGATACAAACAAGTCAAGCGCTTCAT 4479
Qy 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168
Db 4480 AGCATTTGCTTTAGTG-----ACGCTGACCATATTGTTTAGACGA 4518
Qy 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188
Db 4519 GCTGCTAGA-----GATGCGAGTTAAGCAA-----CAATATGAAGCTAAAGCGC 4563
Qy 189 GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluPro 208
Db 4564 GAAATTGAGCAAGCGGAACATCGCACTGATGAAGAAAAACAAGTTGCTTTAAATCAATTA 4623
Qy 209 TyrAlaAsnIleLysAlaAlaLeuGluaspIleThrGlnGluSerAlaMetAspLeuAsn 228
Db 4624 CGGAATAATGAAAAACGTCATTTACAAAACATCGCATCAAGCAATAGCG-----AAT 4674
Qy 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyr 248
Db 4675 AATGATGTGAACGCTGTGAACAAATGCGCATTGCTACATAAAGGTGA----- 4725
Qy 249 AspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAsp 268
Db 4726 -----CAACCTCATATTGTAATTAAGCCT 4749
Qy 269 LeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGlyAlaAsp 288
Db 4750 -----GAAGCACAACAAGCAATAAAGCAAGTGCAGAAAAATCAAGTAGAATCAATAAAA 4803
Qy 289 AspLysAlaPheThrThrVal-----AlaAspGluValProLeuLeuIleGlyAspVal 306
Db 4804 GATACACCACATGCACAGTTGATGAATTAGATGAAGCAATCAATTAATAGCCACACA 4863
Qy 307 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
Db 4864 CTCAAACAAGCGCAACAAGAA-----ATAGAAAAATCAAAATCAAGATGCT 4908
Qy 327 GlyTyrPheAspGlyArg---TrpLeuAspArgSerValIleLeuProasp--- 344
```

```
Db 4909 GCTGTTACTGATGTTAGAAATCAACAATCAAGGCAATAGAGCAAAATAAACCTAAAGTA 4968
QY 345 -----AsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAsp 362
Db 4969 AGAGCTAAACGAGCTGGCTTGTAGATGAAGAAATAATAAAAT----- 5016
QY 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382
Db 5017 -----CAACTCGATGCAATCCGAAATACGTTGGATACTACTCAAGATGAA 5061
QY 383 LeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr 402
Db 5062 AGAGATGTTGCTGTTATGACTTTAAATAAAT----- 5094
QY 403 AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMet 422
Db 5095 -----GTAATACAAATTAATAATGACATGACAA-----AACAAA 5130
QY 423 ValAsnThrGluIleValPheProGluArgGluGlnIleGluAsnAspGlnVal----- 440
Db 5131 ACGAATGCAGAAAGTGCAGACTGAGACTGATGGCAACGACCAACATCAAGTGATTTTA 5190
QY 441 -----SerPheGluGlnSerSerSerSerArgThrGluPro--- 452
Db 5191 CCTAAGTTCAAGTTAAACGACGCGCTCAATCTGTTGGTGTAAAGCCGACGCTCAA 5250
QY 453 ---AlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471
Db 5251 AATGCACATCATCGATCAAGCAGATTATCAACTGAAGAAGAA----- 5292
QY 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491
Db 5293 -----AGACTAGCTGCTAAACATTTAGTAGACAACAGCA 5325
QY 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeu 511
Db 5326 CTTAATCAGGCTATTGATCAGATCAATCATGCA-----GATAAGACTGCC 5370
QY 512 AlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaVal 531
Db 5371 CAAGTTAATCAAGATAGTATTAATGCTCAAAATATTATTCAAAATTAACCCGCGACA 5430
QY 532 SerAlaValAlaArgAlaIle----- 538
Db 5431 ACAGTTAAGCAACAGCAGATACAAACAAATTCGCTACAAAATAAAATTAATTTA 5490
QY 539 -----LeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
Db 5491 ATTAAGCAAAATACGAAGCGACAGATGAGAACAAATATTGCAATAGCACAAGTTGAA 5550
QY 553 -----ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 5551 AAAGAGTTAATTAAGCTAAACAACTAGTGTAGTGCAGTACTAATGCAGATGTG--- 5607
QY 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
Db 5608 -----GCATATTTATTCGATGATGAGAAACGAA----- 5637
QY 587 IleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsn 606
Db 5638 -----ATTCGTGAATCGAA---CCT 5655
QY 607 LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys 626
Db 5656 GTTATTACAGAAGGCGTCTGTCGAGAACAAATTCGACAACTATTATTCACGATATAAAAA 5715
QY 627 GlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAla 646
Db 5716 -----CAAGCAATTGAACGCAATATTCAAGCA 5742
QY 647 ThrLeuGlyTyrClnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666
Db 5743 ACGGTAGAAAGAAAGAAATAGTATATATAGCACAGTTACAAATATTATGACACTGCTATT 5802

US-09-815-242-8815
; Sequence 8815, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8815
; LENGTH: 6561
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6561)
US-09-815-242-8815

Alignment Scores:
Pred. No.: 0.00467 Length: 6561
Score: 135.00 Matches: 154
Percent Similarity: 32.63% Conservative: 110
Best Local Similarity: 19.04% Mismatches: 301
```

```
Query Match: 2.86% Indels: 244
DB: 10 Gaps: 35
US-09-914-168-2 (1-919) x US-09-815-242-8815 (1-6561)

QY 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValPro 43
DB 4144 GTTGCTTCAACAGCATAGCTAAGAAAGAAAGAAAGCACTT----- 4185

QY 44 AlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrPro 63
DB 4186 -----GCAGCTATTGACCACGCT----- 4203

QY 64 GluGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsn----- 76
DB 4204 ---CAAACGAATAGTCAGGTGAATCAAGCAGCAACAATGGTGTATCAGCGATTAAAAATT 4260

QY 77 -----AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGln 92
DB 4261 ATTCACCTGAAACAAAAGTTAAACCCAGCTGCACGTGAATAAATCAAT----- 4308

QY 93 SerProIleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAspMetSerValIle 112
DB 4309 -----CAAAACCGAATCAATACGTGCTAGATTAATCAGATTAAGAAGCAACA 4359

QY 113 GluGluThrThrProLeuSerLeuGlu-----GluLeuPheAlaGlnGluSerThr 129
DB 4360 GCAGAAGAAAGACAAGTAGCAGCTAGATAAAATCAATGAATTTGTAAATCAAGCCATGACA 4419

QY 130 GluMetGlyIleAsn---ProAsnAspTyrIleProGluTyrGlnGlyGlnProAsn 148
DB 4420 GATATACGAATTAATAGAACAAATCAACAAAGTTGATGATCAACAAGCAAGCGCTTGAT 4479

QY 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168
DB 4480 ACCATTGCTTTAGTG-----ACCGCTGACCATATTGTTAGAGCA 4518

QY 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188
DB 4519 GCTGCTAGA-----GATCAGTTTAAGCAA-----CAATATGAAGCTAAAAGCGC 4563

QY 189 GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluPro 208
DB 4564 GAAATTGCAAGCGGAACATCGGCTGATGAAGAAACAAAGTTGCTTTAATCAATTA 4623

QY 209 TyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMetAspLeuAsn 228
DB 4624 GCGAATTAATGAAAACGTCATTACAAAACATCGATCAAGCAATAGCG-----AAT 4674

QY 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyr 248
DB 4675 AATGATGTGAACGCTGTTGAAAACAATGCAATTTGCTACTAAAAGGTGTA----- 4725

QY 249 AspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAsp 268
DB 4726 -----CAACCTCATATTGTAATTAAGCCT 4749

QY 269 LeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGlyAlaAsp 288
DB 4750 -----GAAGCACACAACCAATAAAGCAAGTCAGAGAAATCAAGTCAATCAATAAA 4803

QY 289 AspLysAlaPheThrThrVal-----AlaAspGluValProLeuLeuIleGlyAspVal 306
DB 4804 GATACACCACATGCAACAGCTTGATGAATTAGATGAAGCGCAATCAATTAATAGCGCACAC 4863

QY 307 PheHisHisGlyLysThrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
DB 4864 CTCAAACAAAGCCACAAGAA-----ATAGAAAATCAAAATCAAGATGCT 4908

QY 327 GlyTyrPheAspGlyArg---TrpLeuAspArgSerValAspValIleLeuProAsp--- 344
DB 4909 GCTGTTACTGATGTTAGAAATCAACAATCAAGGCAATAGACCAATAAAACCTTAAGTA 4968
```

```
Db 5803 GGACAAATTGATCAAGATCGTAGC-----AATGCACAAGTTGCATAAAACA 5847
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 5848 GCATCATTAATCTA-----CAAACAATACAT 5874
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 5875 GATTTAGATGATACCTATTAA-----AAGCCAGATGCTGAA----- 5913
Qy 727 AlaValHisLysThrValAlaAsnAsnLeuValAsnPro-----MetArgGlyTyr 743
Db 5914 -----AAACGATTAAATGATCTGCGCGCTGCTGCTTTAGTGCAAAATAT 5964
Qy 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 5965 CCAAAA-----GTAAATATCGTARTAAAGGCT 5991
Qy 764 IleAlaArgAlaGlyIleSerGlyVal 772
Db 5992 GATGCATTAAAGCTATAACTGCTTTA 6018

RESULT 12
US-09-765-272-117
; Sequence 117, Application US/09765272
; Patent No. US200200615451
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-09-765-272-117

Alignment Scores:
Pred. No.: 0.00198 Length: 3121
Score: 134.00 Matches: 190
Percent Similarity: 32.04% Conservative: 124
Best Local Similarity: 19.39% Mismatches: 318
Query Match: 2.83% Indels: 348
DB: 10 Gaps: 53
```

```
US-09-914-168-2 (1-919) x US-09-765-272-117 (1-3121)
Qy 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIle-----IleAsnHis 41
Db 155 TTGACCCAGCCAGATTTCCTGCTCTATATAATAGTCAGCTTCTATCGGAGTCGGGCAACAT 214
Qy 42 ValPro-----AlaHisAspThrAlaIleAsnGlnAlaLys 53
Db 215 TTACACAGAGCCCTCGAAATCGAAGGTTATCAATATATATGTTATATCAAACTAACAA 274
Qy 54 AlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaAla 73
Db 275 CAGGATATACAGAGCTTTCAAGGACA-----GTTGATGGGAAATACTCTGCTCAA 325
Qy 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspAsp----- 91
Db 326 AGAGATAGTCAACCAAACTCTACAAAAAATCAGATAGTTCAGTTCAGCTGATTTAGAA 385
Qy 92 -----GlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeu 107
Db 386 TGGAAACCAAGCAGACAGGGAAGGTTAGTTTACAAGGTGAAGCATCAGGGGATGATGCACTT 445
Qy 108 AspMetSerValIleGluGluThrThrProLeuSerLeuGluLeuPheAlaGlnGlu 127
Db 446 -----TCAGAAAAATCTCTATAGCAGCAGACAATCTATCTCTCTAATGAT 490
Qy 128 SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro 147
Db 491 TCATTCCCAAGTCAAGTTGAGCAGAAT-----CCGATCACAAGGAGGAATCT--- 538
Qy 148 AsnSerGluValValProThrLeuGluProGluLysProGlyLeuIleLysArg 167
Db 539 -----GTAGTTCGACCAACAGTG---CCAGACAAGGAATCTCTGTCTGCT 583
Qy 168 LeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPhe 187
Db 583 ----- 583
Qy 188 TyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGlu 207
Db 584 ---ACAACGGTCGAGAGTCGGGAAGAGGATGTCGCCACGACAATATCGACCAGAG 640
Qy 208 ProTyrAlaAsnIleLysAlaLeuGluAspIle---ThrGlnGluSerAlaMetAsp 226
Db 641 -----TATAAACTTCATTGGAAACCAAGGACGCAAGACCCGCTCATGAG 688
Qy 227 LeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGly 246
Db 688 ----- 688
Qy 247 TyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIle 266
Db 689 -----GGTGAAGCC---CGATGTCGT 706
Qy 267 HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly---Glu 285
Db 707 GAAGACTTA-----CCAGTCTACACT-----AAGCCACTAGAAACCAAGGTACACAA 754
Qy 286 GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAsp 305
Db 755 GGACCCGGACATGAAGGTGAAGCTGCAGTTCGCGGAGGAACACAGCTTACACAAACCG 814
Qy 306 ValPhe-----HisHisGlyLysTyrGluThrLysLysAsn 317
Db 815 TTAGCAACGAAAGCGCAGCAGAGCAGTCTATGAGGCAAGACTACAGTCCGCCAAGAG 874
Qy 318 LeuIleGlu-----AsnAlaSerAlaGluHisGlyTyr 328
Db 875 ACTCTAGATACACCGAACCGGTACGACAAAGGCACACAAAGAACCCCAACAT----- 928
Qy 329 PheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAsp 348
Db ----- 348
```

Db 929 ---GAGGCG-----GAACGSCAGTAGAAGAAGACTTCG----- 961  
Qy 349 ValSerLeuIleTyrAspThrGlyThrGluTyrArgPheAspGluValValPhePheThr 368  
Db 962 ---GCTTTAGAGCTCACTACAGNAATAGAACGGAATCCAGAATATT----- 1006  
Qy 369 IleAspProLysThrAsnGlnLeuThrAspProAspLysLeuProValLysArgGlu 388  
Db 1007 ---CCTTATACAACAGAGAATAATCAGGATCCACACTTCTGAAAAATCGTCGTAAG 1060  
Qy 389 LeuLeuGluGlnLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArg 408  
Db 1061 ATTGAACGACAA-----GGCAACGA----- 1081  
Qy 409 AlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsn-----MetValAsn 424  
Db 1082 -----GGGACACGTCACAAATTCAAATATGAAGACTACATCGTAAT 1120  
Qy 425 ThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln 444  
Db 1121 GGTAAATGTCGTA-----GAAACT 1138  
Qy 445 SerSerSerArgThrGluProAlaGlnValAspGlu-----SerThrLeu 460  
Db 1139 AAGAAGTGTCAACGAAGTCAAGTACCTCCGTCACCAAGACTCGTTAAAGTAGGAACACTT 1198  
Qy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIle 480  
Db 1199 GTGAAGCTTAAACCTACAGTAGAATAATACA-----AACTTAACAAGATT 1243  
Qy 481 GluPheSerAlaSer-----AsnLeuIleGlnAspLysLeuAsnLeuVal 495  
Db 1244 GAGAACAAAAAATCTATAACTGTAACTTAACCTTAATAGACACTACCTCAGCATATGTT 1303  
Qy 496 AlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeuAlaIleAsnHis 515  
Db 1304 TCTCAAAAACGCAA-----GTTTTCCT 1327  
Qy 516 AspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAla 535  
Db 1328 GGAGAC-----AAGCTAGTTAAA 1345  
Qy 536 ArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAla 555  
Db 1346 GAGGTGGATATAGAAAATCTCGCCAAAGAGCAAGTAATA----- 1384  
Qy 556 LeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrVal 575  
Db 1385 -----TCAGGTTTAGATTACTACACACCGTATACAGTTAAACACACCTA 1429  
Qy 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThr 595  
Db 1430 -----ACTTATAATTTGGTGAATAATAGAGAA 1459  
Qy 596 GlyThrArgLeuValThr-----LysPheGluHisAsnLeuIleAsnArgAspGly 612  
Db 1460 AATACTGAACATCAACTCAAGATTTCCAAATTAGATATAGAAAATAGAG----- 1510  
Qy 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla 632  
Db 1511 -----ATTAAGATATTGATTAGTAGAATATATACGGT 1543  
Qy 633 ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln 652  
Db 1544 AAAGAA-----ATGATCGTTATCGTAGATATTAAAGTCTAAGT--- 1582  
Qy 653 GluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGlu 672  
Db 1583 -----GAAGCGCGGACTGATACGGCTAAA----- 1606  
Qy 673 IleSerArgSerIleIleGlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyrArg 692  
Db 1607 -----TACTTTGTAAGAGTGAAA 1624

Qy 693 LeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPhe 712  
Db 1625 TCAGATCGCTTCAAGAAATGTAC-----CTACCTGTAAATCT 1663  
Qy 713 ValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal 732  
Db 1664 ATTACA---GAAAATACGGATGGAACGTATAAAGTAGCGGTAGCCGTTGATCAACTGTCT 1720  
Qy 733 AlaAspAsnLeuValAsnProMetArgTyrArgGlnArgTyrSerLeuGluValGly 752  
Db 1721 GAAGAAGT-----ACAGACGGTTACAAAGATGATTACACATTTACTGTAGCT 1768  
Qy 753 SerSer-----GlyLeuValSerAspAlaAsnMetAlaIleAla----- 765  
Db 1769 AATCTAAACAGAGCAACCCAGAGTTTACACATCCITTAAACAGCTGGTAACGCCATG 1828  
Qy 766 ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsn-----Ala 779  
Db 1829 CAAAGCAATCTGTCTGTCTATACATTTGGCTTCAGATATGACCGCAGATGAGGTGAGC 1888  
Qy 780 TyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSer 799  
Db 1889 TTAGCGCATAGACAGCAAGTTATCTCACAGGTGCATTTACAGGAGCTTGTGCGTTCT 1948  
Qy 800 AspAsnPhe---AsnHisValProTyrArgLeuArg-----PhePheAla 813  
Db 1949 GATGAACAAAATCGTATGCCATTTATGATTTGAAGAAACCATTTATTGATACATTAAT 2008  
Qy 814 GlyGly-----AspGlnSerIleArgGlyTyrAlaHisAspSer----- 826  
Db 2009 GGTCTACAGTTAGAGATTTGGATATTAAACTCTTTCTGCTGATAGATAAAGAAATGTC 2068  
Qy 827 -----LeuSerProIleSerAspLysGlyTyr 835  
Db 2069 GCACCGCTGGCGAAGGACGCAATAGCCGCAATATTATATGTTGCGTAGAGAGGAAAA 2128  
Qy 836 LeuThrGlyGlyGlnValLeu-----AlaValGlyThrAlaGluTyr 849  
Db 2129 ATCTCAGGTGCGAATCTGTTGCGGATTAGTAGCGAGCGCAACAATAACAGTGATGAA 2188  
Qy 850 AsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr 869  
Db 2189 AACAGCTCGTTTACAGGGAACATT-----ATCGCAAAATCACCAG 2227  
Qy 870 AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGly-----ValArgTrp 887  
Db 2228 GACAGTAAT---AAAAATGATAC---GGAGGAATAGTAGGTAAATATACAGGAAAT 2278  
Qy 888 AlaSerProValGlyGlnValArgValAspValAlaThrGlyValLysGluGlyAsn 907  
Db 2279 AGTTGACAGAGTTAATAAGTTAGGTAGATGCCCTTAATCTCTACTAATGCACGCAATAAT 2338

## RESULT 13

US-10-160-758-6  
; Sequence 6, Application US/10160758  
; Publication No. US20030036076A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: CADS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-089C  
; CURRENT APPLICATION NUMBER: US/10/160,758  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 14536

; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-160-758-6

Alignment Scores:

Pred. No.:	0.0211	Length:	14536
Score:	133.50	Matches:	199
Percent Similarity:	31.69%	Conservative:	162
Best Local Similarity:	17.47%	Mismatches:	393
Query Match:	2.82%	Indels:	385
DB:	9	Gaps:	50

US-09-914-168-2 (1-919) x US-10-160-758-6 (1-14536)

QY	52	AlaLysAlaGlyAsnProProValLeuThrProGluGln-----IleGlnAlaArg	69
DB	6446	IITCGGGTGGAGAACGCCATCTCCAGAGTGAGGAAGAGGTACTGTGCACCTGTGCAGA	6505
QY	70	LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPhe	89
DB	6506	AATAAATCCAACCCTGTTTCAGAGTCTTATTACAAGTCAGAGTACCTGAAATATC	6565
QY	90	AspAspGlnSerProIleSerArgIleGlyGlnSerProPro-----LeuGlyLeuAsp	108
DB	6566	ACCCTCTATACCCCAATTTCTCCACACCCAGGCCGCGGAGTCCAGAGGACTCCCGCTCATC	6625
QY	109	MetSerValIleGluGlu-----	114
DB	6626	TACACATTTGTGGAGAGAACCCCTTGATGCTGTTCCACCTGACTTCACAGACTGGTGTC	6685
QY	115	-----ThrThrProLeuSerLeuGlu-----GluLeuPheAlaGlnGlu	127
DB	6686	CTAACAGTAACAGGCCCTTTGGACTATGAGTCCAAGACCAACATGTGTTCCACAGTCAGA	6745
QY	128	SerThrGluMetGlyIleAsnPro---AsnAspTyrIleProGluTyrGlnGlyGluGln	146
DB	6746	GCCAGCGATACAGTCTCGGGTCTATTTCTGAAGCCACAGTGGGAAGTCCTAGTGGAGGAT	6805
QY	147	ProAsnSerGluValValProProThrLeuGlu-----	158
DB	6806	GTCAATGATAAC-----CCTCCCACTTTTTCCCATTGGTCTATACCACCTCCATC	6866
QY	159	-----ProGluLysProGlyLeuIleLysArgLeuTyrAla-----	170
DB	6857	TCAGAAGCGTGCCTGCTCAGACCCCTGTGATCCAACTGTGGCTCTGCACAGGACTCA	6916
QY	171	-----ArgLeuPheAsnAspGlyValAsnLysValProArg	182
DB	6917	GGCGGAACCGTGACGCTCTTATCATGATGTGGAGGATGGCTCAGAT-----	6964
QY	183	LeuLysAlaLysPheTyrGln---SerSerGlnSerGlyGluThrSerAlaIleGlySer	201
DB	6965	---GTTTCCAAGTCTTCCAGATCAATGGGACACAGGGAGAGATGTCACAGTTCAGAA	7021
QY	202	SerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLaLeuGluAsp-----	218
DB	7022	CTGGATTATGAAGCCCAACACACTTCATGTGNAAGTCAGGCCCATGGATAAGAGAT	7081
QY	219	-----IleThrGlnGluSerAlaMet-----AspLeuAsnGlySerIle	231
DB	7082	CCCCCACTCACTGGTGAACCCCTTGTGGTGTCAATGTCTCATATCAATGACACCCC	7141
QY	232	ProArgLeuArgGlnThrAlaLeuValAlaAlaAlaArgAlaValGlyTyrTyrAspIleAsp	251
DB	7142	CCAGATTTCAGACCAACCTCAATATGAAGCCAAATGTCTAGT-----GAA	7183
QY	252	SerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGlu	271
DB	7184	CTGGCAACCTGTGGACACCTGGTCTTTAAGTCCAGGCTATTGACCCCTGACACGACAGAC	7243
QY	272	ProValTyrIleAspTyrArgAlaValAlaGluValArgGlyGlu-----	285
DB	7244	ACCTCCCGCTTGGATACCTGATTTCTTTCGCAATCAGGACAGGCACTCTTCATTAAAC	7303

QY	285	-----
Db	7304	AGCTCATCGGGAATAATTTCTATGTTCAACCTTGTGAAAAAGCACCTGGACTCTTCTTTAC 7363
QY	286	-----GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu 301
Db	7364	ANTTTGAGGTAGTGCTCTGATGGAGTCTTCGAGCAACT-----GTGCCTGTG 7414
QY	302	LeuIleGlyAspVal-----PheHisGlyLysThrGlu 313
Db	7415	TACATCAACACTACAATGCCAACAGTACAGCCAGAGTTCACGACACCTTTATCAG 7474
QY	314	ThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyPheAspGlyArgTrp 333
Db	7475	GCAAGA-----TTAGCAGAGAATGAAATGGTTGGAACCAAGGTGATTGCTGACCC 7528
QY	334	LeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTy 353
Db	7529	ATAGACAAAAGATAGTGT-----CCCTATGGCACTATAGATTATATCATCATCAAT 7579
QY	354	AspThrGlyThrGlnTyArgPheAspGluValValPhePheThrIleAspProLysThr 373
Db	7580	AAACTAGCAAGTGAAG-----TTCTCCATTAAGCCC---AAT 7615
QY	374	AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeu 393
Db	7616	GGCCAGATTGCCACT-----CTGCACAAACTG 7642
QY	394	LeuThrValAsnMetGlyGluAlaTyAsnLeuGlnAlaValArgAlaLeuSerAsnAsp 413
Db	7643	GATCGGAAAAATTCAACAGAG-----AGAGTCATTGCTATTAAAGTTCATGGCTCGGGAT 7696
QY	413	-----
Db	7697	GGAGGAGAACAGTAGCCTTCTCCAGGTGAAGATCATCCTCACAGATGAAATTCACAA 7756
QY	414	-----LeuIleAlaThrArgTyPheAsnMetValAsnThr----- 425
Db	7757	CCCCACAGTTCAAAGCATCTGAGTACACAGATGCATTCATCAATCCAATGTCAGTAAAG 7816
QY	426	-----GluIleValPheProGluArgGluGlnIleGlnAsnAspGluValSer 441
Db	7817	TCCTCCGTTATCCAGGTGTGGCCTATGATGAGATGAAGGTCAAGGTCCGACGATGTCACC 7876
QY	442	PheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461
Db	7877	TAC-----TCAGTGAACCCAGAGACCTAGTTAAAGATGTCAATTGAAATTAAC 7924
QY	462	ProValIleGluThrValGluLeuThrAspGlyIleLeu---MetAspIleSerProfile 480
Db	7925	CCAGTCACTGGTGTGTCAGGTGAAGACAGCCCTGGTGGGATTTGGAAATTCAGACCCCT 7984
QY	481	GluPheSerAla-----SerAsnLeuIleGlnAsp 490
Db	7985	GACTTCTTCATCAAGCCCAAGATGAGGCCCTCTCTCACTGCAACTCTCTGCTCCCACTA 8044
QY	491	LysLeuAsnLeuValAlaAlaLysAlaArg-----HisLeu 502
Db	8045	CGACTTCAGTGGTTCTCTAAAAGATGTCCTTACCAGAAATTTTTCTGAACCTTTTGATACT 8104
QY	503	TyrAspMetProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSer 522
Db	8105	TTCTCTGCACCTGAGAC-----CTTCAGAGGGGTCTGAAATTTGGATTTGTAAGACA 8158
QY	523	IleLeuGlyArgIleSerAspAlaValSerAlaValAlaAlaArgAlaIleLeuProAspGlu 542
Db	8159	GTGCAGCTCAA--GATCCAGTCATCTACAGTCTAGTGGCGGGCACTACACCTTGACAGC 8215
QY	543	SerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrPro 562
Db	8216	AACAAAGATGGTCTCTCTCCCTAGACCCAGACAGAGGGGTCAATAAAGGTGAGGAGGCC 8275

```
Qy 563 AlaAsp-----ValTyrGln----- 567
Db 8276 ATGACCACCAATCCACCAAAATTGTACCAAGATTGATGTGATGGACCATTTGCCTTCAGAAC 8335
Qy 567 ----- 567
Db 8336 ACTGATGTGGTGTCTTGGTCTCTGTCAACATCCAAATGGGAGACGTCATATGACAATAGG 8395
Qy 568 -----SerLysLysValProLeu----- 573
Db 8396 CCTGTATTGAGCGTGATCCATATAAGCGTCTCTCACTGAGATATGCCAGTGGGGACC 8455
Qy 574 -----TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu 589
Db 8456 TCAGTCATTCAAGTGACTGCGCAATTGACAAGGACACTGGGAGAGATGGCCAGGTGAGCTAC 8515
Qy 590 GlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsn 609
Db 8516 AGGCTGTCTGCAGACCCCTGGTAGCAATGTC-----CATGAGCTCTTTGCC 8560
Qy 610 ArgAsp----- 611
Db 8561 ATTGACAGTGAGAGTGGTGGATCACCACACTCCAGGAACCTTGACTGTGAGACCTGCCAG 8620
Qy 612 -----GlyTyrGlnAlaGlyAlaGluLeuArgLeuSer----- 622
Db 8621 ACTTATCATTTTCATGTGTGGTGCCTATGACCAACGACAGACCATCCAGCTATCCTCTCAG 8680
Qy 622 ----- 622
Db 8681 GCCCTGTTTCAGGTCTCCATTACAGATGAGAAATGCTCCCGATTGCTTCTGAA 8740
Qy 623 -----GluAspLysLysGlyValLysLeuTyrAlaThr---Lys 634
Db 8741 GAGTACAGAGGATCTGTGGTTGAGAACAGTGAAGCTCGGCAACTGGTGGCACTCTAAG 8800
Qy 635 ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlu--- 653
Db 8801 ACCCTGGATGCTGACATTTCTTGACAGAACAGGCAAGCTGCTATCATCACAGAGGGA 8860
Qy 654 -----ValPheGlyHisSerThrAsnGlyPheAsp-----LeuSerThr 666
Db 8861 GACCCCTGGCCAGTGTGGCATGACGCAAGTTGGAGATGATGGAGGATTTCTCTCAAG 8920
Qy 667 ArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr 686
Db 8921 AAGACCTGGACCGCAGCATACAGCCAAG----- 8950
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 8951 TACTTGTCTCAGAGTCACAGCATCTGTATGGCAAGTTCAGGCT---TCGGTCACTGTGGAG 9007
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 9008 ATCTTTGCTGGACGTCATGATGATAACAGCCACAGTGTTCACAGCTTCTGTATACTGGC 9067
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 9068 AAGGTTTCATGAAGATGTA-----TTTCCAGGACAC 9097
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArg 766
Db 9098 TTCATTTTGAAGGTTTCTGCCACAGACTTGGACACTGATCAATGCTCAGATC----- 9151
Qy 767 AlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHis 786
Db 9152 -----ACATATTCCTG-----CATGCCCTGGG---GCGCAT 9181
Qy 787 GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSer----- 799
Db 9182 GAATTCAAGCTGGATCTCATACAGGGGAGCTGACACCATCTACTGCTAGACCCGAGAA 9241
Qy 800 -----AspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGln 817
```

```
Db 9242 AGGAAGGATGTGTTCAACCTTGTT-----GCCAAGCGCAGGATGGAGGTGGCCGA 9292
Qy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
Db 9293 TCGTGGCAGGAGACATCACCTCCATGTGGAGGATGTGAATGACAATGCC----- 9343
Qy 838 GlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeu 857
Db 9344 -----CCGCGGTTCTTCCCCCAGGCACCTGT 9367
Qy 858 ArgLeuAlaValPheGlyAsp-----IleGlyAsnAlaTyrAspLysGlyPheThrAsn 875
Db 9368 GCTGTGCTGTCTTCGACACACACACATGAGAGCCCTGTGTGGCTGTAGTATTTGCCCGG 9427
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerPro-----ValGly 892
Db 9428 GATCCCCAGCAAGCGGCAATGCCAGGTGTTTACTCTCTGCGCGGATTCACCCGAAGGC 9487
Qy 893 GlnValArgValAspValAlaThrGlyValLysGluGlyGluGlyAsnProIleLysLeu 911
Db 9488 CACTTTTCCATCGACGCCACCAACCGGGGTGATCCGCTGGAAGAACCCGCTGCAGGTC 9544
```

## RESULT 14

US-10-160-758-7

; Sequence 7, Application US/10160758  
; Publication No. US20030036076A1

## GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: CADS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-089C  
; CURRENT APPLICATION NUMBER: US/10/160,758  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 14536  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-160-758-7

## Alignment Scores:

Pred. No.:	0.0211	Length:	14536
Score:	133.50	Matches:	199
Percent Similarity:	31.69%	Conservative:	162
Best Local Similarity:	17.47%	Mismatches:	393
Query Match:	2.82%	Indels:	385
DB:	9	Gaps:	50

US-09-914-168-2 (1-919) x US-10-160-758-7 (1-14536)

Qy 52 AlaLysAlaGlyAsnProProValLeuLeuThrProGluGln-----IleGlnAlaArg 69  
Db 6446 GCTCGGGATGGAGAACCCCATCTCCAGAGTGAAGAGAGTACTTGTCTCACTGTGAGA 6505  
Qy 70 LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValIleAsnPhe 89  
Db 6506 AATAAATCCAAACCCACTGTTTTCAGAGTCCTTATTACAAGTCAGAGTACCTGAAATATC 6565  
Qy 90 AspAspGlnSerProIleSerArgIleGlyGlnSerProPro---LeuGlyLeuAsp 108  
Db 6566 ACCCTCTATACCCCAATTTCTCCACACCCAGCCCGGAGTCCAGAGGAGTCCGGCTCATC 6625  
Qy 109 MetSerValIleGluGlu----- 114  
Db 6626 TACAACATGTGGAGGAGAAGAACCTTGATGCTGTTCACCACTGACTTCAAGACTGGTCTC 6685

Qy	115	-----ThrThrProLeuSerLeuGlu-----GluLeuPheAlaGlnGlu	127	Db	7643	GATCGGAAATTCACAGAG-----AGAGTCATTGCTATTAAAGGTCATGGCTCGGGAT	7699
Db	6686	CTAACAGTAAACAGGGCCTTTGGAGCTATGATCCAAAGCAAAACATCTGTTCACAGTCACA	6745	Qy	413	-----	413
Qy	128	SerThrGluMetGlyIleAsnPro---AsnAspTyrIleProGluTyrGlnGlyGluGln	146	Db	7697	GGAGGAGGAAGAGTAGCTTCTGCACGGTGAAGATCATCTCACAGATCAAAATGACAAC	7756
Db	6746	GCACGGATACAGCTCTGGGTCATTTCTGAAGCCACAGTGGAAAGTCTAGTGGAGGAT	6805	Qy	414	-----LeuIleAlaThrArgTyrPheAsnMetValAsnThr-----	425
Qy	147	ProAsnSerGluValValValProProThrLeuGlu-----	158	Db	7757	CCCCACAGTTCAAAGCATCTCAGTACAGATATCCATTCAATGCCAANTGCAGTAAAGAC	7816
Db	6806	GTCAATGATAAC-----CTCCCACTTTTCCCAATTTGGTCTATACCACCTTCCATC	6856	Qy	426	-----GluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSer	441
Qy	159	-----ProGluLysProGlyLeuIleLysArgLeuTyrAla-----	170	Db	7817	TCCTCGGTTATCCAGGTGTGCCTATGATGCAGATGAAGGTGACAACGCAGATGTCCACC	7876
Db	6857	TCAGAAGGCTTCCTGCTCAGACCCTGTGATCCCACTGTTCGGCTTCACACAGGACTCA	6916	Qy	442	PheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu	461
Qy	171	-----ArgLeuPheAsnAspGlyValAsnLysValProArg	182	Db	7877	TAC-----TCAGTGAACCCAGGAGCCTAGTTAAAGATGTCAATTGAAATTAAC	7924
Db	6917	GGCGGAACCGTAGCTCTTATCAGATTTGGAGGATGGCTCAGAT-----	6964	Qy	462	ProValIleGluThrValGluLeuThrAspGlyIleLeu---MetAspIleSerProIle	480
Qy	183	LeuLysAlaLysPheTyrGln---SerSerGlnSerGlyGluThrSerAlaIleGlySer	201	Db	7925	CCAGTCACTGGTGTGCTCAAGGTGAAAGACAGCCTGGTGGGATTTGAAATCAGACCCCTT	7984
Db	6965	---GTTTCCAGTCTTCCAGCATCAATGGGAGCACAGGGAGATGTCACAGTTCACAGAA	7021	Qy	481	GluPheSerAla-----SerAsnLeuIleGlnAsp	490
Qy	202	SerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuLysVal	218	Db	7985	GACTTCTTCATCAAAAGCCCAAGATGGAGGCCCTCTCTCACTGGAACACTCTCTGTGGCAGTA	8044
Db	7022	CTGGATTATGAAGCCCAACACACTTTCATGTGAAGTTCAGGGCCATGGATAAGAGAT	7081	Qy	491	LysLeuAsnLeuValAlaAlaLysAlaArg-----HisLeu	502
Qy	219	-----IleThrGlnGluSerAlaMet-----AspLeuAsnGlySerIle	231	Db	8045	CGACTTCAGGTGGTTCCTAAAAAGTAGTCTTACCAGAAATTTTCTCAACCTTTGTATATCT	8104
Db	7082	CCCCACACTCATGGTGAACCCCTGTGGTGTGTCAATGTGTCTGATATCAATGACAACCC	7141	Qy	503	TyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSer	522
Qy	232	ProArgLeuArgGlnThrAlaLeuValAlaAlaAlaArgAlaValGlyTyrTyrAspIleAsp	251	Db	8105	TTCTCTGCACCTGAAGAC-----CTCCAGAGGGTCTGAAATTTGGATTTGTTAAAGCA	8158
Db	7142	CCAGAGTTCACAGCACTCAATATCAACCAATGTCAGT-----GAA	7183	Qy	523	IleLeuGlyArgIleSerAspAlaValSerAlaValAlaAlaArgAlaIleLeuProAspGlu	542
Qy	252	LeuSerIleIleArgAsnSerIleGlyValValAspValIleIleHisAspLeuGlyGlu	271	Db	8159	GTGGCAGCTCAA---GATCCAGTCACTACAGTCTAGTTCGGGGCAGCTACACCTGAGACC	8215
Db	7184	CTGGCAACCTGTGCACACCTGGTCTTAAAGTCCAGGCTATTGACCCCTGCACAGCAGAC	7243	Qy	543	SerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrPro	562
Qy	272	ProValTyrIleAspTyrArgAlaValAlaGluValArgGlyGlu-----	285	Db	8216	AACAAGGATGGTCTTCTCCCTAGACCCAGACACAGGGGTCTATAAAGGTGAGGAAGCCC	8275
Db	7244	ACCTCCGCCCTGGAGTACCTGATCTTCTGGCAATCAGGACAGCAGCCTTCTTCATTAAAC	7303	Qy	563	AlaAsp-----ValTyrGln-----	567
Qy	285	-----	285	Db	8276	ATGGACCCAGAAATCCACCAAAATGTATCCAGATGTGATGGCCACATTTGCCTTCAGAAC	8335
Db	7304	AGCTCATCGGGAATAATTTCTATGTTCAACCTTTGCAAAAAGCAGCCTGGAGCTCTCTTAC	7363	Qy	567	-----	567
Qy	286	-----GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu	301	Db	8336	ACTGATGTGGTCTCTTGGTCTCTGTCAACATCCAAGTGGGAGAGCGTCAATGACAATAGG	8395
Db	7364	AATTTGAGGTAGTGCTCTGATGGAGTCTTCGAGCAACT-----GTGCCCTGTG	7414	Qy	568	-----SerLysLysValProLeu-----	573
Qy	302	LeuIleGlyAspVal-----PheHisHisGlyLysTyrGlu	313	Db	8396	CCTGATTTGAGGTGATCCATATTAAGCTCTCTCACTGAGATATGCCAGTGGGAGCC	8455
Db	7415	TACATCAACACTACAANAATGCCAACAGTACAGCCAGAGTTCAGCAGCAGCCTTTATGAG	7474	Qy	574	-----TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu	589
Qy	314	ThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyr	333	Db	8456	TCAGTCAATCAAGTGACTGCCATTGACAAGCAGACACTGGGAGAGATGGCCAGGTGAGCTAC	8515
Db	7475	GCAGAA-----TTAGCAGAGAATGCAATGGTGGAAACCAAGGTGATTGTTGCTAGCC	7528	Qy	590	GlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsn	609
Qy	334	LeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyr	353	Db	8516	AGGCTGTCTCGACAGCCCTGGTAGCAATGTC-----CATGAGCTCTTTTGGC	8560
Db	7529	ATAGACAAAGATAGTGGT-----CCCTATGGCACTATAGATTATATCATCATCAAT	7579	Qy	610	ArgAsp-----	611
Qy	354	AspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThr	373	Db	8561	ATTGACAGCTGAGAGTGGTTGGATCACCACACTCCAGAACTTGTACTGTGAGACCTGGCAG	8620
Db	7580	AAACTAGCAAGTGAAGAAG-----TTCTCCATAAACCCTC---AAT	7615	Qy	612	-----GlyTyrGlnAlaGlyAlaGluLeuArgLeuSer-----	622
Qy	374	AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeu	393	Db	8621	ACTTATCATTTTTCATGTGGTGGCCTCATGACCCAGCAGACCATCCAGCTATCTCTCAG	8680
Db	7616	GGCAGATTCGCCACT-----CTGCAGAAACTG	7642	Qy	622	-----	622
Qy	394	LeuThrValAs					



Db 8681 GCCCTGGTTTCAGGCTCTCCATTACAGATGAGATGAGAAATGCTCCCGATTTGCTTCTTGAA 8740  
Qy 623 -----GluAspLysLysGlyValLysLeuTyrAlaThr---Lys 634  
Db 8741 GAGTACAGAGGATCTGGTTGAGAACAGTGGCCGAACTGGTGGCGACTCTAAG 8800  
Qy 635 ProLeuSerHisProLeuAsnAspGlnLeuAlaThrLeuGlyTyrGlnGlnGlu--- 653  
Db 8801 ACCCTGGATGCTGATCTTCTGAGCAGAACAGGCGAGGTCACCTGCTACATCACAGAGGA 8860  
Qy 654 -----ValPheGlyHisSerThrAsnGlyPheAsp-----LeuSerThr 666  
Db 8861 GACCCCTGGCCGAGCTTTGGCATGACCAAGTTGGAGATGGAGGAGATTTCCTCAGG 8920  
Qy 667 ArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyTyrAsnArgThr 686  
Db 8921 AAGACCTGGCCGCGAGCATACGCCAAG----- 8950  
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706  
Db 8951 TACTTGTCTGGAGCTCAATGATACAGCCACAGTTCAGAGCT---TCGGTCACTGTGGAG 9007  
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726  
Db 9008 ATCTTTGCTGGAGCTCAATGATACAGCCACAGTGTTCACAGCTTCTCTATCTACTGCG 9067  
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746  
Db 9068 AAGTTTCATGAAGATGTA-----TTTCCAGGACAC 9097  
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArg 766  
Db 9098 TTCAATTTGAAGGTTTCTGCCACAGACTTGGACACTGATACCAATGCTCAGATC----- 9151  
Qy 767 AlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHis 786  
Db 9152 -----ACATATCTCTG-----CATGGCCCTGGG---CGCAT 9181  
Qy 787 GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSer----- 799  
Db 9182 GAATTCAGAGCTGGATCTCTATACAGGGAGGCTGACACACTCAGTCCCTAGACCGAGAA 9241  
Qy 800 -----AspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyClyAspGln 817  
Db 9242 AGGAAGGATGTGTTCAACCTTGTT-----GCCAAGGCGACGGATGGAGGTGGCCGA 9292  
Qy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837  
Db 9293 TCGTCCAGGCGAGCATCACCTCCATGTGGAGGATGGAATGACAATGCC----- 9343  
Qy 838 GlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeu 857  
Db 9344 -----CCGCGGTTCTTCCCGAGCCACTGT 9367  
Qy 858 ArgLeuAlaValPheGlyAsp-----IleGlyAsnAlaTyrAspLysPheThrAsn 875  
Db 9368 GCTGTGGCTGTCTTCGAAACACACACAGTGAAGACCCCTGTGGCTGTAGTATTGCCCGG 9427  
Qy 876 AspThrLysIleGlyValGlyValArgTyrAlaSerPro-----ValGly 892  
Db 9428 GATCCCGACCAAGCGCCCAATGCCAGGTGGTTTACTCTCTCCCGGATTCACGCCAAGGC 9487  
Qy 893 GlnValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeu 911  
Db 9488 CACTTTTCCATCGAGCCACCGCGGGGTGATCCGCTGGAAAGCCGCTGCAGGTC 9544

## RESULT 15

US-09-815-242-8512

; Sequence 8512, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8512  
; LENGTH: 17388  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(17388)  
US-09-815-242-8512

## Alignment Scores:

Pred. No.: 0.0309 Length: 17388  
Score: 133.00 Matches: 165  
Percent Similarity: 31.24% Conservative: 144  
Best Local Similarity: 16.68% Mismatches: 348  
Query Match: 2.81% Indels: 332  
DB: 10 Gaps: 45

US-09-914-168-2 (1-919) x US-09-815-242-8512 (1-17388)

Qy 51 GlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArg--- 69  
Db 2374 CGGTAAACCACTACATACACCAATTAACACGACGCTGAGCAATTAAGAGGTACAGCA 2433  
Qy 70 -----LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu 84  
Db 2434 TTACAAAAAGTACCTGTTAATATTTCTGGAATACCGTTCGATCCATCACCATTGTTAT 2493  
Qy 85 AspValValAsnPheAspAspGlnSerProIleSerArgIle-----GlyGluGln 101  
Db 2494 TTAGTT-----CGGCCAACAAATCAAACTACGAATGCTGTAGTGAG 2535  
Qy 102 SerProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 121  
Db 2536 GCAGATCAAAATACCATCTGTTATACGATCTGGACTGGTACCTGATGGGTCAT 2595  
Qy 122 GluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyr----- 138  
Db 2596 AACACAATTACT-----ATACACCAAAAGATTACGTAGTATT 2634  
Qy 139 IleProGluTyrGlnGlyGluGln----- 146  
Db 2635 ATACCACCCAGTA---GGTAAACAAATTAGAGCAGTAGTTATTATAATAAAGTAGTTGCA 2691  
Qy 147 -----ProAsnSerGluValValValProProThrLeuGluProGluLysProGly 163  
Db 2692 TCTAATATGAGTAATGCTGTTACTATTTTGGCCAGATGACATTCACCAACA----- 2742

[illegible][illegible]

```
Db 4405 CAAATACGGAAATFACTACTGCTATTAAACATTTATTTGTACCAAGTGATTTAGTG 4464
Qy 828 SerProileSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAla- 843
Db 4465 AATCCAACAATTTTCAGTAGGTAANTACACTAATCATCATCAAGTGTCTCAGGTGAACAATTT 4524
Qy 843 ----- 843
Db 4525 ACAAACTATTATACAGCGAATGATAAATTTTGGTGTGCAATCGGTAACGTACCAACTACA 4584
Qy 844 -----ValGlyThrAlaGluTyrAsnTyrGluPheMet- 854
Db 4585 TCACAAATTACAGGTACTGTTGATAATAACCATCAACATGTTTCAGCAACAGCACAAT 4644
Qy 855 -----LysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAla 868
Db 4645 GTGACATCAGCACTAGTAAACAACAATCAATTTATTAGCAACTGATACAGTGGTAATACA 4704
Qy 869 TyrAspLysGlyPheThrAsnAspThr-----LysIleGlyAlaGly 882
Db 4705 GCTACAACCTTCATTC-----AATGTAACAGTGAACACCTTTGCGTGATAAATATATCGAGTTGGT 4761
Qy 883 ValGlyValArgTyrAlaSerProVal 891
Db 4762 ACTTCATCAACGGGTGCTAATCTGCTG 4788

RESULT 16
US-09-815-242-9166
; Sequence 9166, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9166
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1488)
US-09-815-242-9166

Alignment Scores:
Pred. No.: 0.000949 Length: 1488
Score: 132.50 Matches:
```

```
Percent Similarity: 34.47% Conservative: 79
Best Local Similarity: 21.05% Mismatches: 197
Query Match: 2.80% Indels: 189
DB: 10 Gaps: 33

US-09-914-168-2 (1-919) x US-09-815-242-9166 (1-1488)

Qy 148 AsnSerGluValValProThrLeuGluProGluLysProGlyLeuIleLysArg 167
Db 4 TCATCTAAGGTTATGTT-----ACAATTTTCGGTGGAGAGACCTGGCTAAACGC 57
Qy 168 ---LeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla 186
Db 58 AAGCTCTACCTTCCTCTTTT-----AGA 81
Qy 187 PheTyrGlnSerSerGln---SerGlyGluThrSerAlaIleGlySerSerHisGlnLys 205
Db 82 CTATATAATCCGCAATCTTTCCAGCACTTTGCGGTATTTGGAACCTGCCCGTAGA--- 138
Qy 206 ThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGluSerAlaMet 225
Db 139 -----CCTTGGAGT-----AAGGAATATTTTGAATCTGTAGTTGTCGAGTCCATCCTT 186
Qy 226 AsnLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaVal 245
Db 187 GATTTGGCAGATAGTACCGCAGCAAGCCCAAGAAATTTGCT-----AGCCAC 231
Qy 246 GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle 265
Db 232 TTCTACTATCAA----- 243
Qy 266 IleHisAspLeuGlyGluProVal---TyrIleAspTyrArgAlaValGluVal 282
Db 244 AGCCATGATGTCAATGATTGCGAACATTTATTTGCGTCAATTTACAACTCAGCTT 303
Qy 283 -----ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValPro 300
Db 304 AATGAAAAATACCAAGCTGAACACAAATAAGCTCTTC-----TTCTTGTCTATGCACT 357
Qy 301 LeuLeuIleGlyAspValPheHisGlyLysTyrGlyThrLysLysAsnLeuIleGlu 320
Db 358 CAGTCTTTGGACCATTTGCCAACACCTCAATCTGAA-----AACATTGTC--- 405
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspVal 340
Db 406 -----GATGCAAAAGGTTTGGAGCGC----- 426
Qy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGln 358
Db 427 -----TTGATCGTTGAAAAACCATTTGTCACAGAT 456
Qy 359 Tyr-----ArgPheAspGluValValPhe 366
Db 457 TACGCAACTGCAACAAGTGAATGACCACTCTAGCAACATTTGACGAACAACAAAT 516
Qy 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386
Db 517 TTCCGTATCGACCATTA-----CTTGGT 540
Qy 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406
Db 541 AAGGAATGATCCAAAGCATCTTTGCGTCTGCTTTGCA-----AACTTGATTTT 591
Qy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlu 426
Db 592 GAAACAGTTTGGACAAGCATTTTATCCAC-----AATGTTCAA 630
Qy 427 IleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer 446
Db 631 ATTACCTTTGCGAGCGCTTGGGTCTAGAACACCTGGTGGTACTATGACCAATCCGCT 690
Qy 447 SerSerArgThrGluProAlaGlnValAspGluSerThrLeuLeuProValIleGluThr 466
Db --- 111
```

```
Db 691 GCCCTCCGTGAC-----ATGCTCCAAACCACACTCTA----- 723
Qy 467 ValGluLeuThrAspGlyLeuMetAspIleSerProIleGluPheSerAlaSerAsn 486
Db 724 ---CAACTCTTTCCTCTCGCTCCATGAC---AAACCGCAAGCTTCACAAAGACGAG 777
Qy 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
Db 778 ATTCTGCTGAAAGATTAAAGTCTTT-----AAAACTCTATCAT---CCA 822
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
Db 823 ACTGATGAAGAACTCAAGAACTTTATCCGTGGCAATACCCCTCT-----GGTAAG 876
Qy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
Db 877 ATT---GATGCGCATGAATACATCTCTATCTGCTAGCGAGCCAAAT----- 918
Qy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 919 -----GTGAATCCAGAATCAACAACACTGAAACCTTT 948
Qy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgGlyGln 586
Db 949 ACATCTGGTGC-----TTCTTTGTAGACAGCGATGATTCCTG----- 987
Qy 587 IleGlyLeuGlyTyrPglySerAspThrGlyThrArgLeuValThrLysPheGluHis--- 605
Db 988 ---GGTCTCTCTTTCTTTTCCGTACAGGTAAACGACTGACTGAAAAAGGAACCTCATGTC 1044
Qy 606 AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLys 625
Db 1045 AACATCGTC----- 1053
Qy 626 LysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg 645
Db 1054 -----TTTAAACAATGGATTCTATCTTTGGAGAACCACTTGCT 1092
Qy 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665
Db 1093 CCAATATAT-----TTGACCATCTATATTCAACCAACAGAGGCTTCTCTTAGC 1143
Qy 666 ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArg 685
Db 1144 CTAATGGGAAGCAAGTAGGAGAAGATTAACTTGGCTCCTAAC----- 1188
Qy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrp 705
Db 1189 -----TCACTTGATTACCGTACAGATGCGACTGCAACTGGTCTTCCAGAACCATAC 1242
Qy 706 GlnAspLeuProValAspPheValAsn 714
Db 1243 GAAATTTGATTTATGATGCTCTAAAT 1269

RESULT 17
US-09-815-242-7037
Sequence 7037, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Cart, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
```

```

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7037
; LENGTH: 2571
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2571)
US-09-815-242-7037

Alignment Scores:
Pred. No.: 0.00211 Length: 2571
Score: 132.50 Matches: 109
Percent Similarity: 36.35% Conservative: 96
Best Local Similarity: 19.33% Mismatches: 212
Query Match: 2.80% Indels: 147
DB: 10 Gaps: 25

US-09-914-168-2 (1-919) x US-09-815-242-7037 (1-2571)
Qy 28 AlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThr 47
Db 343 GCTTTAGAAGAAGAGGAAACGATCAGCGATATTTTGAAGAAAGTGGCGTGAAGAAAGAA 402
Qy 48 AlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGln 67
Db 403 CAAATTTCCCAAGCT-----ATTGAGCAGCATTTAGA 432
Qy 68 AlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 87
Db 433 GCGGACAAAACGCTGAACGATCAAAATGCAGAAAGAAAGCAGACAAAGCGCTTGAAATAAT 492
Qy 88 AsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeu 107
Db 493 ACGATT---GATTTAACCGCTCGTCGAGAAAGTGGCAAACTTGATCTGTATTTGGCGCT 549
Qy 108 AspMetSerValIleGluGluThrThrProLeuSerLeuGluLeuPheAlaGlnGlu 127
Db 550 GAT-----GAAGAAATTT----- 561
Qy 128 SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro 147
Db 562 -----CGTCGAGCCATTAAGTATTACAACGCTCGTACCAA 597
Qy 148 AsnSerGluValValProProThrLeuGluPro-----GluLysProGlyLeuIle 165
Db 598 AATAACCTGTGTTAATTGGT-----GAACGAGGTAGGAGAAACCGCGATTTGTC 648
Qy 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArg---LeuLys 184
Db 649 GAAGGCTTGGCAGCAGCGCATCGTAAAC-----GGCGAAGTGGCAGAGGTTTCAAA 699
Qy 185 AlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGln 204
Db 700 AATAACCGTGTCTTTTATTAGATATGGGG-----CGGTTGATTGCTGGTGGCAATAAT 753
Qy 205 LysThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGlnSerAla 224
Db 754 CGTGTGTAATTTGAAGAACGCTTTAAAGCAGCTACTCAATGAACCTTTGAAAGAACAGAGGT 813
```

```

Qy 535 AlaAraGalLalIe 538
    ||| |||||
Db 1741 GCAAACGCGATT 1752

RESULT 18
US-09-815-242-4761
; Sequence 4761, Application US/09815242
; Patent No. US20020061369A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: EUI7A.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4761
; LENGTH: 7434
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4761

Alignment Scores:
Pred. No.: 0.0112 Length: 7434
Score: 132.00 Matches: 176
Percent Similarity: 32.22% Conservative: 123
Best Local Similarity: 18.97% Mismatches: 326
Query Match: 2.79% Indels: 304
DB: 10 Gaps: 40

US-09-914-168-2 (1-919) x US-09-815-242-4761 (1-7434)

Qy 2 SerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuA
    :||| ||||| ||| :|||
Db 1423 ACTAAGAAGTTTATTATCCAGATAAATCT-----

Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnAsnAsnProAlaAsnI
    ||| ||| :||| :|||
Db 1456 AAATTATCATATAAGTTAATGTTGCGAATATCATACACCTAAATAATA

Qy 39 -----IleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaL
    :||| :||| :||| :|||
Db 1516 GAAAAATTTACATATCTACTGCTTCAGATGTTCTTAATTAATAATGCGC

Qy 57 ProProValLeuLeuThr-----
    ||| ||| |||||
Db 1567 CCAGAAGTTACACTTAACTCGAGATCCATTTTCAGTACGGTGTGAATGAC

Qy 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAla-----
    :||| :||| :||| :|||
Db 1627 TTGCAACCAACGAATGAATCAAGTTGATTAATAGTCTATTACACACACACGA

```

Qy	78	-----LysProGlnSerGlnAlaLeuAspValValAsnPheAsp-	-----	90
Db	1687	GAATACAATAAAGCTTAAACAACAAGCAGATAGTATTTTAAATGAAGATGCGAATCATGTT	:    :    :    :    :	1746
Qy	91	-----AspGlnSerProIle-	:	95
Db	1747	AAAGCTCAAATCGTCACTCAAGCGGATATGATGGTTAGTAACTAAATTTACAAGCT	:	1806
Qy	96	SerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGluGluThr	:    :    :    :    :	115
Db	1807	GCATTAATTGATAATCAAGCAGCAATTCGTGAATTAGATACTAAAGCTCAAGAAAGGTT	:    :    :    :	1866
Qy	116	ThrProLeuSerLeuGluLeuPheAlaGlnGluSer-	:    :    :    :    :	130
Db	1867	ACAGCAGCACAAAGATAAAAGTTACGCAAGATGAAGTTGCAGCAGCTTGTAACCTAAA	:    :    :    :    :	1926
Qy	131	MetGlyIleAsnProAsnAspTyrIleProGluTyr-	:    :    :    :    :	144
Db	1927	ATTAACAATGATAAAATAATGCATCCGAGAAATATATAACAACACTACAGCAAGGT	:    :    :    :    :	1986
Qy	145	-----GluGlnProAsn-	:    :    :    :    :	156
Db	1987	GTCACAACTAAAAAGATAATGATCGCAGTGTAGAACCAAGATGTGATTACACCAACA	:    :    :    :    :	2046
Qy	157	LeuGluProGlu--LysProGlyLeuIleLysArgLeuTyrAlaArg-	:    :    :    :    :	171
Db	2047	GTTAAACCTCAAGCGAAACAAGATATTATCCAAAGCAGTTACAACCTCGTAACACACAAATT	:    :    :    :    :	2106
Qy	172	-----LeuPheAsnAspGlyValAsnLys	:    :    :    :    :	179
Db	2107	AAAAGTCAAATGCATCATTTACAAGATGAAAAGATGTAGCAAAATGATAAAATTCGTAAA	:    :    :    :    :	2166
Qy	180	ValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly-	:    :    :    :    :	194
Db	2167	ATTGAAACAAAGGCAATTAAAGATATTGATGCAGCAACAACAATGCACAAGTAGAAGCC	:    :    :    :    :	2226
Qy	195	--GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLys	:    :    :    :    :	213
Db	2227	ATTAAACAACAGCAATCAATGATATTAACTCAACTACACCTGCTACACAGCTTAAGCA	:    :    :    :    :	2286
Qy	214	AlaAlaLeuGlu-	:    :    :    :    :	217
Db	2287	GCAGCTCTTGAAGAATTTGACGAAGTTGTCAGCACAAATTTGATCAAGCACCTTTAAAT	:    :    :    :    :	2346
Qy	218	---AspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGln	:    :    :    :    :	236
Db	2347	CCTGATACACAATGAAGAAGTAGCGGAA--GCTATTGACGCTATTAAATGCA	:    :    :    :    :	2397
Qy	237	ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle-	:    :    :    :    :	254
Db	2398	GCTAAAGTTTCTGGTGTAAAGCAATTTGAGCGCAACGACTGCACAAGATTTTAGAAGA	:    :    :    :    :	2457
Qy	255	IleArgAsn-	:    :    :    :    :	257
Db	2458	GTTAAACACGAAGAAATCTCAAAAATTTGAAATATTTACTGACTCTACGCAACCAAAATG	:    :    :    :    :	2517
Qy	257	-----	:    :    :    :    :	257
Db	2518	GATGCCCTTAATGAAGTTAAACAAGCTGCCACAGCTAGAAAAGCTCAAAATGCTACAGTT	:    :    :    :    :	2577
Qy	258	-----SerIleGlyGluValAspValIleIleHisAspLeuGlyGlu	:    :    :    :    :	271
Db	2578	TCAATGCCAACAATCAAGAAGTAGCAGAAGCTGATGCAGCAGTAGATGCAGCTCAAAAG	:    :    :    :    :	2637
Qy	272	ProValTyrIleAspTyrArgAlaValGluValArgGlyGlyAlaAspAspLys--	:    :    :    :    :	290
Db	2638	CAAGGTTTACATCACATCCAAGTTGTTAAATCAAAACAAGGAAGTTGCTGATACAAATCA	:    :    :    :    :	2697
Qy	291	-----AlaPheThrValAlaAspGluValProLeuLeuIle	:    :    :    :    :	303
Db	2698	AAAGTATTAGATAAAATCAATGCAATTCACACCAAGCAAGGTTAAACCT--GCA	:    :    :    :    :	2751

Qy	304	GlyAspValPheHisHISgLyTyTyrGluThrLysLysAsnLeuIleGluAsnAlaSer	329
Db	2752	GCTGCTACGGAGTAGAAAACGCGATATATAATACGTAACAAGAAATTTCAAAATAGCAAT	2811
Qy	324	AlaGlu-----HisGlyTyTyrPheAspGlyArgTrpLeuAsp	335
Db	2812	GCTTCAACTACAGAAGAAAAACAAAGCTGCATATACAGAAATTAGATTACTATAAACACAGCAAA	2871
Qy	336	ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyTyrAsn	354
Db	2872	GCAAGAACAAATCTTGATGCTGCAAAATACAAACAGTGATGTAAACAACAGCTAAAGACAAAT	2931
Qy	355	-----ThrGlyThrGlnTyArgPheAsp-----	362
Db	2932	AGTATTGCTGCAATTATCAAGTTCAAGCTGCCACAACCTAAGAAAATCGGATGCAAAAACGC	2991
Qy	363	GluValValPhePheThrIleLeuAspProLysThr-----AsnGlnLeuThr	377
Db	2992	GAAATCGTCAAAAAGCAAGTGAACGTAAACAGCAATTTGAAGCAATCAATGATTCGACT	3051
Qy	378	ThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuThrValAsn	397
Db	3052	ACTGAAGAACAAACAA-----GCAGCGAAGACAAAGTGGATCAAGCAGTAGTTACTGCCA	3105
Qy	398	MetGlyGluAlaTyAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThr	417
Db	3106	AACGCTGATATAGATAAT-----GCTGCAGCAACAAATGATGTGGATTAATGCCA	3153
Qy	418	ArgTyTyrPheAsnMetValAsnThrGluIleValPheProGlu-----	431
Db	3154	AAAACTACAAATCAAGCTACAAATCGCAGCGATTTACACCTGATGCAAAATTTTAAACACGAC	3213
Qy	432	-----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer	446
Db	3214	CGAAACACAGCAATTCAGATAAAGTACAAAGCTCAAGAAACAGCAATTTGATGGAATAATAC	3273
Qy	447	SerSerArgThrGlu-----ProAlaGlnValAspGlnSerThr	459
Db	3274	GGCTCAACAACCTGAAGAAAAAGCAGCTGTAAACAACAAGTTCAACCTGAAACAAACAACA	3333
Qy	460	LeuGluProValIleGlu-----ThrValGlnLeuThrAspGlyIleLeuMet	475
Db	3334	GCTGATCGCCGCAATAGATCGACGACATACAAATGCGGAAGTTGAAGCGGCTAAAAAAGCA	3393
Qy	476	AspIleSerProIleGluPhe-----SerAlaSerAsn	486
Db	3394	GCAATTCGCTAAAATTTGAAGCGATTACGCGACCAACAACAACCTAAAGATAATGCGAAAGAA	3453
Qy	487	LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyTyrAspMetPro	506
Db	3454	GCAATTCGTCGAAGCGGAATGAACTAAACAGCAATCGCTCAACCGCAAGCATTTACT	3513
Qy	507	AspArgValLeuAlaIleHisAspAspGlyValAsnArgSerIleLeuGlyArg	526
Db	3514	GCTCAAGAAATTCAGCGGCTAATGCGGAC-----	3543
Qy	527	IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu	546
Db	3544	GTAGATAATGCTGTGACACAAAGCAATAGCACATTTGAAGCTGCTAATGACTCAAAATGAT	3603
Qy	547	Val-----IleAsp-----	549
Db	3604	GTACACCAACGGAACAGCAGGTGAAATAGTATTGATCAAGTAACCAACCAACAGTTAAT	3663
Qy	550	-----LeuProGluArgThrAlaLeuAlaAsnArgLys-----	560
Db	3664	AAAAAGCAACTGCAGGTAAATGAAATCACAGCAATTTTAAATAACAATTCGAAGAGATT	3723
Qy	561	-----ThrProAlaAspValTyTyrGlnSerLysLysValProLeuTyTyrValPheValAla	578
Db	3724	CAAGCTACCGCAGATGCACAGATGAAGAAAAAAGCAGCGT-----GATGCT	3771
Qy	579	SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyTyrGlySerAspThrGlyThrArg	598

```
Db 3772 GAAGCAATACGTAATGGTAAAGCAAAATCAAGCCATTTCAGCAGCAACTACTAACGCA 3831
      :::::::::::
Qy 599 LeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlu 618
      :::::
Db 3832 CAAGTT-----GATGAGCTAAAGCAATGCAGAA 3861
      :::::::::::
Qy 619 LeuArgLeuSerGlu----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSer 637
      :::::
Db 3862 GCAGCGATTAAATGCGGTAAACACCAAAAGTTGTGAAGAAACAAGCGCTAAAGATGAAAT 3921
      :::::
Qy 638 HisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis 657
      :::::::::::
Db 3922 -----GATCAATTACAAGCAACG-----CAACAATAATTATCATCAATAAT 3960
      :::::::::::
Qy 658 SerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIle 677
      :::::
Db 3961 -----GATCAGAACGCTACACAGAGAAAGAAAGAGC-AGCTATTCAACA 4004
      :::::
Qy 678 IleGlnAsnGlyGlyTrpAsnArg 685
      :::::
Db 4005 ATTAGCAACAGCAGTTACAGAGCG 4028
      :::::
```

## RESULT 19

```
US-09-815-242-8869
; Sequence 8869, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIORITY FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8869
; LENGTH: 7437
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(7437)
US-09-815-242-8869

Alignment Scores:
Pred. No.: 0.0112 Length: 7437
Score: 132.00 Matches: 176
Percent Similarity: 32.22% Conservative: 123
Best Local Similarity: 18.97% Mismatches: 326
Query Match: 2.79% Indels: 304
DB: 1.0 Gaps: 40
```

```
US-09-914-168-2 (1-919) x US-09-815-242-8869 (1-7437)
Qy 2 SerLysProValPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeu 21
      :::::
Db 1423 ACTAAGAAGTTTATTCCAGATAAATCT-----TTA 1455
      :::::::::::
Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIle----- 38
      :::::
Db 1456 AAATTATCATATAAAGTTAATGTTGCGAATATCGGATACACCTAAAAATATTGATTTAAT 1515
      :::::
Qy 39 -----IleAsnHisValProAlaHisaspThrAlaIleAsnGlnAlaLysAlaGlyAsn 56
      :::::
Db 1516 GAAAAATTTAAACATATCGTACTGCTTCAGATGTTGTAATTAATTAATGCGCAA----- 1566
      :::::::::::
Qy 57 ProProValLeuLeuThr-----ProGluGln 65
      :::::
Db 1567 CCAGAGTTTACACTAACCTGCAGATCCATTTTCAGTAGCGGTTGAATGAACAAAGATGCG 1626
      :::::
Qy 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAla----- 77
      :::::
Db 1627 TTGCAACAACAAGTAAACTCACAAGTTGATAATAGTCTATTACACAACAGCATCAATTGCA 1686
      :::::::::::
Qy 78 -----LysProGlnSerGlnAlaLeuAspValValAsnPheAsp----- 90
      :::::
Db 1687 GAATACAATAAACTTAAACAACAAGCAGATACTATTTTAAATGAAGATGCGAATCATGTT 1746
      :::::
Qy 91 -----AspGlnSerProIle----- 95
      :::::
Db 1747 AAACTGCAATTCGTGCATCTCAAGCGGATATTGATGTTTAGTAACATAAATTACAAGCT 1806
      :::::::::::
Qy 96 SerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGluGluThr 115
      :::::
Db 1807 GCATTAATTGATAATCAAGCAGCAATTCGTGAATTAGATACTATAAGCTCAACAAAGGTT 1866
      :::::::::::
Qy 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSer-----ThrGlu 130
      :::::
Db 1867 ACAGCAGCACACAAGTAAAAAAGTTACGCAAGATGAAGTTGCGAGCACTTGTACTATAA 1926
      :::::::::::
Qy 131 MetGlyIleAsnProAsnAspTyrIleProGluTyr-----GlnGly 144
      :::::
Db 1927 ATTAACAATGATAAAAATAATGCAATCGCAGAAATTAATAAACAACACTACACCAAGGT 1986
      :::::::::::
Qy 145 -----GluGlnProAsn-----SerGluValValValProThr 156
      :::::
Db 1987 GTCACAACCTGAAAAAGATAATGGTATCGCAGTGTAGAACAAAGATGTGATACACCAACA 2046
      :::::::::::
Qy 157 LeuGluProGlu---LysProGlyLeuIleLysArgLeuTyrAlaArg----- 171
      :::::
Db 2047 GTTAAACCTCAAGCGAAACAAGATATTATCCAAGCAGTTACAACTCGTAAACAACAATA 2106
      :::::::::::
Qy 172 -----LeuPheAsnAspGlyValAsnLys 179
      :::::
Db 2107 AAAAAGCTCAATGCATATTACAGATGAAAAAGATGTAGCAAAATGATAAAATTTGTTAA 2166
      :::::::::::
Qy 180 ValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly----- 194
      :::::
Db 2167 ATTGAACAACAAGGCAATTAAGATATTGATGCGACCAACAACAATAATGCAAGTAGAAGCC 2226
      :::::::::::
Qy 195 ---GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLys 213
      :::::
Db 2227 ATTAACAACAAGCAATCAATGATATTAACTCAACACTACACCTGCTACACAGCTAAAGCA 2286
      :::::::::::
Qy 214 AlaAlaLeuGlu----- 217
      :::::
Db 2287 GCAGCTCTTGAGAGATTTGACGAGAGTTGTTCAAGCACAAATTTGATCAAGCACCCTTTAA 2346
      :::::::::::
Qy 218 ---AspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGln 236
      :::::
Db 2347 CCTGATACAACAATAATGAAGAGTAGCGGAA-----GCTATTGAAGCTATTAAATGCA 2397
      :::::::::::
Qy 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle-----AspLeuSerIle 254
      :::::
```

Db 2398 GCTAAAGCTTTCTGGTGTAAAGCAATTGAAGCGACACAGCTGCACAAAGATTTAGAAAGA 2457  
Qy 255 IleArgAsn----- 257  
Db 2458 GTTAAAAACGAAGAAATCTCAAAAATTGAAATATTACTGACTCTACGCAACAAAAATG 2517  
Qy 257 ----- 257  
Db 2518 GATGCCCTAATGAAGTTAAACAAGCTGCAACAGCTAGAAAAGCTCAAAAATGCTACAGTT 2577  
Qy 258 -----SerIleGlyValValAspValIleIleHisAspLeuGlyGlu 271  
Db 2578 TCAANTGCACAAATGAAGAGTAGCAGAGCTGATGCCAGCAGTAGTAGTCAGCTCAAAAG 2637  
Qy 272 ProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys--- 290  
Db 2638 CAAGTTTACATGACATCCCAAGTTGTTAAATCAAAACAGCAAGTTGCTGATACAAAATCA 2697  
Qy 291 -----AlaPheThrThrValAlaAspGluValProLeuLeuIle 303  
Db 2698 AAAGTATTAGATAAAATCAATGCAATTTCAAAACACAGCAAAAAGTTAAACCT-----GCA 2751  
Qy 304 GlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSer 323  
Db 2752 GCTGATACGGAAGTAGAAAACGCTATATACACGTAAACAGAAATTTCAAAATAGCAAT 2811  
Qy 324 AlaGlu-----HisGlyTyrPheAspGlyArgTyrLeuAsp 335  
Db 2812 GCTTCAACTACAGAAGAAAACAGCTGCATATACAGAAATAGATATAAAAAGCAAGAA 2871  
Qy 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp--- 354  
Db 2872 GCAAGAACAAATCTGATGCTGCAAAATACAAACAGTAGTATACACAGCAGCTAAAGACAAT 2931  
Qy 355 -----ThrGlyThrGlnTyrArgPheAsp--- 362  
Db 2932 AGTATTGCTCAATTAATCAAGTTCCAGCTGCCACAACTAAGAAATCGGATGCAAAAGCG 2991  
Qy 363 GluValValPhePheThrIleAspProLysThr-----AsnGlnLeuThr 377  
Db 2992 GAAATCGCTCAAAAAGCAAGTGAACGTAAACAGCAATTTGAAGCAATGAATGATTCGACT 3051  
Qy 378 ThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsn 397  
Db 3052 ACTGAAGAACAAACA-----GCAGCGAAAGACAAAGTGGATCAAGCAGTAGTTACTGCA 3105  
Qy 398 MetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThr 417  
Db 3106 AACGCTGATATAGATAT-----GCTGCACCAACAACATGATGTGGATAATGCA 3153  
Qy 418 ArgTyrPheAsnMetValAsnThrGluIleValPheProGlu----- 431  
Db 3154 AAAACTACAATGAAGCTACAACTCGCAGCCATTACACTGATGCAAAATGTTAAACCCAGCA 3213  
Qy 432 -----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer 446  
Db 3214 GCAAAACAACAATGCGAGATAAGTACAAAGCTCAAGAAACACCAATGATGGAATAAAC 3273  
Qy 447 SerSerArgThrGlu-----ProAlaGlnValAspGluSerThr 459  
Db 3274 GGCTCAACAACCTGAGAAAAGCAGCTGCTAAACAACAAGTTCAACTGAAAAACAACA 3333  
Qy 460 LeuGluProValIleGlu-----ThrValGluLeuThrAspGlyIleLeuMet 475  
Db 3334 GCTGATGCCCAATAGATGACGACATACAAATGCGGAAGTTCAAGCGGCTAAAAAGCA 3393  
Qy 476 AspIleSerProIleGluPhe-----SerAlaSerAsn 486  
Db 3394 GCAATGCTAAAATGAAGCGATTTCAGCCAGCAACAACTAAAGATAATGCGAAAGAA 3453  
Qy 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506  
Db 3454 GCAATGCTACGAAGCGAATGAACGTAAACACAGCAATCGCTCAACAGCAAGACATTACT 3513

Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526  
Db 3514 GCTGAAGAAATTCAGCGGCTAATGGCGAC----- 3543  
Qy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546  
Db 3544 GTAGATAATGCTGTGACACAAGCAAAATAGCAACATTTGAAGCTCTAATAGTCAAAATGAT 3603  
Qy 547 Val-----IleAsp----- 549  
Db 3604 GTAGACCAACGAAACGACAGCTGAAAATAGTATTGATCAAGTAACCAACAGTTAAT 3663  
Qy 550 -----LeuProGluArgThrAlaLeuAlaAsnArgLys----- 560  
Db 3664 AAAAAAGCAACTGCACGTAAATGAAATCACACGAATTTTAAATTAACAAATTTGCAAGAGATT 3723  
Qy 561 -----ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAla 578  
Db 3724 CAAGCTACGCCAGATCAACAGATGAAGAAAACAAACAGAGCT-----GATGCT 3771  
Qy 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArg 598  
Db 3772 GAAGCAAAATCTGAAATGCTAAAGCAAAATCAAGCCATTTTCAGCAGCACTACTAAGCA 3831  
Qy 599 LeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlu 618  
Db 3832 CAAGTT-----GATGAAGCTAAAGCAAAATGCAGAA 3861  
Qy 619 LeuArgLeuSerGlu---AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSer 637  
Db 3862 GCAGCGATTAAATCGGTACACCAAAAGTTGTGAAGAAACAAGCGGCTAAAGATGAAATT 3921  
Qy 638 HisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHis 657  
Db 3922 -----GATCAATTTACAGCAACG-----CAAACAATGTTATCAATAAT 3960  
Qy 658 SerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIle 677  
Db 3961 -----GATCAGAAGCTTACAAACAGAAAGAAAGCAAGC-AGCTATTCAACA 4004  
Qy 678 IleGlnAsnGlyGlyTyrAsnArg 685  
Db 4005 ATTAGCAACAGCAGTTACAGACGC 4028  
RESULT 20  
US-09-815-242-4648  
; Sequence 4648, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27



; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4648  
; LENGTH: 6077  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-4648

## Alignment Scores:

Pred. No.: 0.00938 Length: 6077  
Score: 131.50 Matches: 184  
Percent Similarity: 33.76% Conservative: 135  
Best Local Similarity: 19.47% Mismatches: 345  
Query Match: 2.78% Indels: 281  
DB: 10 Gaps: 48

US-09-914-168-2 (1-919) x US-09-815-242-4648 (1-6077)

```
Qy 48 AlaIleAsnGlnAlaLys-----AlaGlyAsnProProValLeuLeuThrProGlu 64
Db 942 GCAGTGAATGACGCTAAACAAACAACTTACATGGTGTGATCAAAATTTAGCTCAAGATAAGCAA 1001

Qy 65 GlnIleGlnAlaAArgLeu---AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAla 83
Db 1002 CGTGCACAGAAACCTTAAATAACTTGTCTAACTTGAATACA--CCACAGCTCAAGCA 1058

Qy 84 LeuAsp-----ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGln 101
Db 1059 CTTGAAATCAAAATCAATCAATGACCAACTCGTGGCGAAGTAGCACAAAAATTAAGTGGAG 1118

Qy 102 SerProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 121
Db 1119 GCACAAGCATTAAACCAAGCAATGAAGCATTTACGTAAT-----AGCAITCAA 1166

Qy 122 GluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGlu 141
Db 1167 GAT-----CAACAGCAACACAGACGGGT-----AGCAAGTTTATCAATGAA 1208

Qy 142 -----TyrGlnGlyGluGlnProAsnSerGluValValValPro 154
Db 1209 GATAAGCCACAAAAAGATGCATACCAGCAGCAGTACAAACATGCCAAAAGATTGATTAAT 1268

Qy 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174
Db 1269 CAACTAGCAATCCA-----ACGCTTGATAAGCACAAAGTTGCAACAATTGACA 1316

Qy 175 AspGlyValAsnLysValPro-----ArgLeuLysAlaLysPheTyrGlnSerSer 191
Db 1317 CAAGGTGTTAACCAAGCTAAAGATAACCTACATGGTGTGATCAAAACATTCGACACGATAA 1376

Qy 192 GlnSerGlyGluThr-----SerAlaIleGlySerSerHisGlnLys 205
Db 1377 CAACATGCGAGTTACTGATTAAATCAATTAATAGTTTAAATAAATCCGCAAGCTCAAGCT 1436

Qy 206 ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlu----- 222
Db 1437 CTTGAAAGTCAAAATAACACCGCAACTCGTGTGATGAAGTAGCGCAAAAATTAAGTACGAA 1496

Qy 223 ---SerAlaMetAsp-----LeuAsnGlySerIleProArgLeuArgGln 236
Db 1497 GCACAAGCGCTTGATCAAGCAATGCAAGCATTACGTAAATAGTATTCAAGATCAACAACAA 1556

Qy 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg 256
Db 1557 ACAGAA---TCTGGTAGCAAGTTTATCAATGAAGATAAGCCCAAAAAAGATGCATACCAA 1613

Qy 257 AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAsp 276
Db 1614 GCAGCAGTACAAACATGCCAAAGATTTAATTAACCAACAGGTAATCCAACA---CTCGAT 1670
```

```
Qy 277 TyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296
Db 1671 AAATCACAGTTGAACAATTA-----ACACAGCAGTACAACTGCAAAA 1715

Qy 297 AspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys 316
Db 1716 GATAACCTA-----CATGGTGATCAAAAACTTGCTCGT 1748

Qy 317 AsnLeuIleGlu-----AsnAlaSerAlaGluHisGlyTyrPheAspGly 331
Db 1749 GATCAACAACAGCAGTAAACGACTGTTAATGCATTACCAAACTTAAATCATCGGCAACAA 1808

Qy 332 ArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu 351
Db 1809 CAAGCTTTAACTGATGCTATAAATGCA--CGCCCTACAGAACACA--CAGGTTGCACAA 1862

Qy 352 IleTyrAspThrGlyThrGlnTyr-----ArgPheAsp 362
Db 1863 CATGTTCAAACTGCTACTGAACCTTGATCAGCGATGGAACACATTGAAAAAATAAGTTGAT 1922

Qy 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382
Db 1923 CAAGTGAATACAGATAAGGCTCAACCAATTACACTGAAGCGTCAACTGAT----- 1973

Qy 383 LeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal----- 396
Db 1974 -----AAAAAGAGCAGTAGATCAAGCGTTTACAAGCTGCAGAAAGCATTACAGAT 2024

Qy 397 ---AsnMetGlyGluAlaTyrAsnLeuGlnAlaVal---ArgAlaLeuSerAsnAspLeu 414
Db 2025 CCACTAATGTTGTTCAAAATGCGGAATAAGACGCTGTAGAACACGATTAAGTAACTAAG 2078

Qy 415 IleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGln 434
Db 2079 CTTCAAGAAAAAGTAAATGAGTTAAATGGTAATGAGAGAGTGCCTGAAGCTAAAGCACAA 2138

Qy 435 IleGlnAsn-----AspGlnValSerPheGluGlnSerSerSerSerArgThrGluPro 452
Db 2139 CGGAAACAACAACTATTGACCAATTTAGCACATTTAAATGCTGTGATCAAAATTCGAACCTGCTAAA 2198

Qy 453 AlaGlnValAspGluSerThr---LeuGluProValIleGluThrValGluLeuThrAsp 471
Db 2199 CAAAACTTTGATCAAGCAGCAAACTTCAACCAATCGTGAATATTAGTAGAT----- 2249

Qy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491
Db 2249 ----- 2249

Qy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeu 511
Db 2250 -----CAAGCAACGCAATTTGAATCAATCTATGATCAATTCGAACAA 2291

Qy 512 AlaIleAsnHisAspAspGlyValAsnArgSerIle-----LeuGlyArgIleSerAsp 529
Db 2292 CGAGTCAATGAACATGCTAAACGTTGAGCAAACTAGATATTACACGCAAGCAGATTCAGAT 2351

Qy 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn----- 545
Db 2352 AAGCAAAATGCTTATTAACAAAGCTATTGCA-----GAAGCTGAAAATGTATTAAACAA 2405

Qy 546 -----GluVal 547
Db 2406 AATTCGAATAAGCAACAGTGGATCAAGCAGCTTCAAAACATTTTAAATGCAAAACAAAGCA 2465

Qy 548 IleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGln 567
Db 2466 TTAATGTTGATGAACGTTGATAGCTTGTCTAAACAAATAAGTAAACATGACATCGACCAA 2525

Qy 568 SerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIle 587
Db 2526 TTGAATGCAATTA----- 2537
```

```
Qy 588 GlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeu 607
Db 2538 -----ACCAAT 2543
Qy 608 IleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGly 627
Db 2544 GCTCAACAAGATGATTTAAAGGT-----CGCATCGATCAA----- 2579
Qy 628 ValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn----- 641
Db 2580 -----TCACAGATTAAATCAATCAAAATCCAAACAAATTTGTA 2612
Qy 642 -----AspGlnLeuArgAlaThrLeuGlyTyrGln 651
Db 2613 GATCAGGCTAAGGCACCTAATCGTGAATGATCAATG-----TCA 2654
Qy 652 GlnGluValPheGlyHis-----SerThrAsnGlyPheAspLeuSer 665
Db 2655 CAAGAAATCTCTGCAATGAAGGACGACGAAAGGTAGCAGCACTATGTCAATGCAGAT 2714
Qy 666 ThrArgThrLeuGlu-----HisGluIleSerArgSerIleIleGln 679
Db 2715 ACACAACTTAAACAAGTGTATGATGAACGGTGTATGAAGCAACAGCACTTGATAA 2774
Qy 680 AsnGlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699
Db 2775 TCGACAGTCAAAACTTAATCGCAGACAAAGTTATCAAAATTAATGATGCGAGTCACTGCA 2834
Qy 700 AlaProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro----- 717
Db 2835 GCTAAGAAAGCA-----TTAAATGCTGAAGAAAGACTT 2867
Qy 718 -----SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAsp 734
Db 2868 AATAATCGTAAGTCTGAAGCAATTAACAAGA-----TTGGAT 2903
Qy 735 AsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSer 754
Db 2904 CAATTAACACATCTAAACAATGCTCAAAAGACAATTAGCAATCCAAACAATTAATAATGCT 2963
Qy 755 GlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSer 774
Db 2964 GAAACGCTAAATAAGCATCTCGAGCAATTAATAGAGCG-----ACT 3005
Qy 775 PheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAla 794
Db 3006 AATTAGATAATGCAATGGGTGCA-----GTACAACAATATATTGACGAACACACTT 3059
Qy 795 GlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGly 814
Db 3060 GGTGTTATCAGCAGCACAAATTAATCATCAATGATGATAATTTGAAA-----GCAAT 3113
Qy 815 GlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGly 834
Db 3114 TATGATAATGCAATTCGAATGCGACCATGAG-----TTAGATAAA--- 3155
Qy 835 TyrLeuThrGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMet 854
Db 3156 -----GTGCAAGGTAAATGCAATTCGAAGTGAAGCAGACAGCAATTTGAAACAAATAT 3209
Qy 855 LysAspLeuArgLeuAlaValPheGlyAsp-----IleGlyAsnAlaTyrAspLys--- 871
Db 3210 ATCGATGCTCAAAATGCAATTAATGGAGACCAACAACTTGCAAAATGGCAAGATAAAGCA 3269
Qy 872 ---GlyPheThrAsn 875
Db 3270 AATGCGTTTGTAAAT 3284
```

RESULT 21  
US-09-815-242-8513  
; Sequence 8513, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cart, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8513
; LENGTH: 9477
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; NAME/KEY: CDS
; LOCATION: (1)...(9477)
US-09-815-242-8513
Alignment Scores:
Pred. No.: 0.018 Length: 9477
Score: 131.50 Matches: 184
Percent Similarity: 33.76% Conservative: 135
Best Local Similarity: 19.47% Mismatches: 345
Query Match: 2.78% Indels: 281
DB: 10 Gaps: 48
US-09-914-168-2 (1-919) x US-09-815-242-8513 (1-9477)
Qy 48 AlaIleAsnGlnAlaLys-----AlaGlyAsnProProValLeuLeuThrProGlu 64
Db 4336 GCAGTGAATGACGCTAAACAACTTACATGGTGTATCAAAAATTAGCTCAAGATAAGCAA 4395
Qy 65 GlnIleGlnAlaArgLeu---AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAla 83
Db 4396 CGTCCACACAGAAAGCTTAAATAAATCTGCTTAACCTGAATACA---CCACAACGTTCAAGCA 4452
Qy 84 LeuAsp-----ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGln 101
Db 4453 CTTGAAATCAAAATCAATAATGCAACAACTCGTGGCGAAGTAGCACAAAAATTAACGTAG 4512
Qy 102 SerProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 121
Db 4513 GCACAAGCACTTACCACCAAGCAATGGAAGCATTCAGTAAAT-----AGCATTTCAA 4560
Qy 122 GluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGlu 141
Db 4561 GAT-----CAACAGCAAAACAGACGGGT-----AGCAACTTTTATCAATCAA 4602
Qy 142 -----TyrGlnGlyGluGlnProAsnSerGluValValValPro 154
Db 4603 GATAAGCCACAAAAAGATGTCATACCAAGCAGCAGTACCAAGATGCGCAAGATTTGATTAAT 4662
Qy 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174
```

Db 4663 CAAACTAGCAATCCA- : : : : : ACGCTTGATAAAGCAGCAAGTTGAACAATTGACA 4710  
Qy 175 AspGlyValAsnLysValPro- : : : : : ArgLeuLysAlaLysPheTyrGlnSerSer 191  
Db 4711 CAAGGTGTTTACCAAGCTAAAGATAACCTACATGCTGTATCAAAACTTGCAGACGATAAA 4770  
Qy 192 GlnSerGlyGluTrn- : : : : : SerAlaIleGlySerSerHisGlnLys 205  
Db 4771 CAACATGCAGTTACTGATTAAATCAATTAAATAGTTTAAATAATCCGCAACGCTCAAGCT 4830  
Qy 206 ThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGlu- : : : : : 222  
Db 4831 CTTGAAGTCAAAATAACAACGCAACCTCGTGATGAAGTAGCGCAAAAATTAGCTGAA 4890  
Qy 223 ---SerAlaMetAsp- : : : : : LeuAsnGlySerIleProArgLeuArgGln 236  
Db 4891 GCACAAGCGCTGTGATCAAGCAATGCAACGATTACCTATATGATTATCAAGATCAACAACAA 4950  
Qy 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg 256  
Db 4951 ACAGAA---TCTGGTAGCAAGTTTATCAATGAAGATAAGCCCAAAAGAGATGCATACCAA 5007  
Qy 257 AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAsp 276  
Db 5008 GCAGCAGTACACATGCCAAAGATTTAATTAACCAAAACAGGTAATCCAAACA---CTCGAT 5064  
Qy 277 TyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296  
Db 5065 AAATCACAAGTTGACAATTA- : : : : : ACACAAGCAGTACAACTGCCAANA 5109  
Qy 297 AspGluValProLeuLeuIleGlyAspValPheHisGlyLysThrGluThrLysLys 316  
Db 5110 GATAACCTA- : : : : : CATGGTGATCAAAAACCTTCTCGT 5142  
Qy 317 AsnLeuIleGlu- : : : : : AsnAlaSerAlaGluHisGlyTyrPheAspGly 331  
Db 5143 GATCAACAACAGCAGTAAACGACTGTTAATGCATTACCAAACTTAAATCATCGCGCAACAA 5202  
Qy 332 ArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu 351  
Db 5203 CAAGCTTTAACTGATGCTATAAATGCA---GCGCTACAGNAACA---CAGGTTGCCAACA 5256  
Qy 352 IleTyrAspThrGlyThrGlnTyr- : : : : : ArgPheAsp 362  
Db 5257 CATGTTCAAACTGCTACTGAACCTGATCAGCGATGGAACATTTCAAAAATAAAGTTGAT 5316  
Qy 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382  
Db 5317 CAAGTGAATACAGATAAGGCTCAACAAATTTACACTGAAGCGTCAACTGAT- : : : : : 5367  
Qy 383 LeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal- : : : : : 396  
Db 5368 -----AAAAAGACGCTAGATCAAGCGTTACAGCTGCAGAAAGCATTACAGAT 5418  
Qy 397 ---AsnMetGlyGluAlaTyrAsnLeuGlnAlaVal---ArgAlaLeuSerAsnAspLeu 414  
Db 5419 CCAACTAATGGTTCAAAATCGGAATAAAGACGCTGTAGAACAGCATTAACCTAAG- : : : : : 5472  
Qy 415 IleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGln 434  
Db 5473 CTTCAAGAAAAAGTAAATGAGTTAAATGTAATGAGAGAGTGCCTGCAAGCTAAAGCACAA 5532  
Qy 435 IleGlnAsn- : : : : : AspGlnValSerPheGluGlnSerSerSerArgThrGluPro 452  
Db 5533 GCGAAACAACAACTATTGACCAATTTAGCACATTTAAATGCTGATCAAAATTCGAACCTGCTAAA 5592  
Qy 453 AlaGlnValAspGluSerThr---LeuGluProValIleGluThrValGluLeuThrAsp 471  
Db 5593 CAAAACATTGATCAAGCGAGAACTTCAACCAATCGTGAATTAGTAGAT- : : : : : 5643  
Qy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491

Db 5643 ----- 5643  
Qy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeu 511  
Db 5644 -----CAAGCAACGAANTGAATCAATCTATGGATCAATTTGCAACA 5685  
Qy 512 AlaIleAsnHisAspAspGlyValAsnArgSerIle- : : : : : LeuGlyArgIleSerAsp 529  
Db 5686 GCAGTCAATGAACATGCTTAACGTTGAGCAAACTGTAGATTACACCAAGCAGATTCAGAT 5745  
Qy 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn- : : : : : 545  
Db 5746 AAGCAAAATGCTTATAAAACAAGCTATTGCA- : : : : : GAAGCTGAAAAATGTATTAACAA 5799  
Qy 546 -----GluVal 547  
Db 5800 AATTCCAATAAGCAACAAGTGGATCAACGACTTCAAAACATTTTAAATGCCAANAACGA 5859  
Qy 548 IleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGln 567  
Db 5860 TTAATGCTGATGAACGTGTAGCAGCTTCTAAACAAATGGTAAACATGACATCGACCAA 5919  
Qy 568 SerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIle 587  
Db 5920 TTGAATGCATTA- : : : : : 5931  
Qy 588 GlyLeuGlyTyrGlySerAspThrArgLeuValThrLysPheGluHisAsnLeu 607  
Db 5932 -----AACAA 5937  
Qy 608 IleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGly 627  
Db 5938 GCTCAACAAGATGGATTTAAAGGT- : : : : : CGCATCGATCAA- : : : : : 5973  
Qy 628 ValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn- : : : : : 641  
Db 5974 -----TCACACAGATTTAAATCAAAATCCAAACAAATTTGTA 6006  
Qy 642 -----AspGlnLeuArgAlaThrLeuGlyTyrGln 651  
Db 6007 GATGAGGCTAAGGCCTTAATCGTCAATGGATCAATTG- : : : : : -TCA 6048  
Qy 652 GlnGluValPheGlyHis- : : : : : SerThrAsnGlyPheAspLeuSer 665  
Db 6049 CAAGAAATCTCTGGCAATGAAGGAGCGACGAAAGGTAGCAGCAACTATGTCAATGACAT 6108  
Qy 666 ThrArgThrLeuGlu- : : : : : HisGluIleSerArgSerIleIleGln 679  
Db 6109 ACACAAGTTAAACAAGTGTATGATGAAGCGTTGATAAGCGAAACAAGCAGCTTGATAAA 6168  
Qy 680 AsnGlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699  
Db 6169 TCGACAGGTCAAAACCTTAACCTGCAGAACAGTTATCAAAATTAATGATGCAGCTCACTGCA 6228  
Qy 700 AlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro- : : : : : 717  
Db 6229 GCTAAGAAAGCA- : : : : : TTAATCGTGAAGAAAGACTT 6261  
Qy 718 -----SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAsp 734  
Db 6262 AATAATCGTAAGCTGAAGCATTTACAAGA- : : : : : -TTGGAT 6297  
Qy 735 AsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSer 754  
Db 6298 CAATTTAACACATCTAAACAATGCTCAAGACAAATTAGCAATCCAAACAAATTAATAGCT 6357  
Qy 755 GlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSer 774  
Db 6358 GAAACGCTAAATAAAGCATCTCGAGCAATTAATAGAGCG- : : : : : -ACT 6399  
Qy 775 PheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAla 794  
Db 6400 AAATTTAGATAATGCAATGGGTGCA- : : : : : -GTACAACAATATATTACGACGACGACCTT 6453

QY 795 GlyTyrIleTrpSerAspAsnHisValProTyrArgLeuArgPheAlaGly 814  
Db 6454 GGTGTTATCAGCAGCAAAATATCATCATCCAGATGATAATTGAAA-----GCAAAAT 6507  
QY 815 GlyAsnGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGly 834  
Db 6508 TATGATAATGCAATTCGCAATGCGACACATGAG-----TTAGATAAAA--- 6549  
QY 835 TyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMet 854  
Db 6550 -----GTGCAAGGTAAATGCAATTCGCAAAAGCTGAACGAGCAATTCGCAAAATATT 6603  
QY 855 LysAspLeuArgLeuAlaValPheGlyAsp-----lleGlyAsnAlaTyrAspLys--- 871  
Db 6604 ATCGATGCTCAAAATGCAATTAATGAGACCAAAACCTTGCAAAATGCGCAAAAGATAAGCA 6663  
QY 872 ---GlyPheThrAsn 875  
Db 6664 AATCGGTTTGTGTAAT 6678  
RESULT 22  
US-08-781-986A-59  
; Sequence 59, Application US/08781986A  
; Publication No. US20030054436A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781.986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31096 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-781-986A-59  
Alignment Scores:  
Pred. No.: 0.102 Length: 31096  
Score: 131.50 Matches: 181  
Percent Similarity: 33.45% Conservative: 118  
Best Local Similarity: 20.25% Mismatches: 324  
Query Match: 2.78% Indels: 271  
DB: 7 Gaps: 43  
US-09-914-168-2 (1-919) x US-08-781-986A-59 (1-31096)

QY 31 GlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAlaIleAsn 50  
Db 28877 CAAACTAAACAATCCACCGCTTGATAAGACCAAGTTGCAACAATTGACACAACTGTTAAAC 28936  
QY 51 GlnAlaLys-----AlaGlyAsnProProValLeuLeuThrProGluGlnIleGln 67  
Db 28937 CAAGCTAAACATAACCTTACACGGTGATCAAAAACCTTGACACAGCATAAACAACATGCGGTT 28996  
QY 68 AlaArgLeuAsnAlaAla---GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86  
Db 28997 ACTGATTTAAATCAATTAATGTTTGAAT---AATCCGCAACCTCAAGCACTT----- 29047  
QY 87 ValAsnPheAspAspGlnSerProIleSerArgIle-----GlyGluGlnSerPro 103  
Db 29048 -----GAAAGCCAAATAACAACAGCAGCAACTCGTGCGCAAGTAAAGATAAA 29092  
QY 104 ProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluLeu 123  
Db 29093 AAATTAGCT---GAAAGCAAAAGCGCTTGATCAAGCAATGCAAGCATTCAGTAATAGTATT 29149  
QY 124 PheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGlu----- 141  
Db 29150 CAAGATCAACAACAACAGAGATCTGGT-----AGCAAGTTTATCAATGAAGATAAA 29200  
QY 142 -----TyrGlnGlyGluGlnProAsnSerGluValValVal----- 153  
Db 29201 CCGCAAAAAGATGCTTACCAAGCAGCAGCACTTCAAAATGCAAAAGATTTTAATTAACCAACA 29260  
QY 154 --ProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAla----- 170  
Db 29261 GGTANTCCAACTCGCACAAAATTCACAGTAGAACAAATTCACACAAAGCAGTAACAACACTGCA 29320  
QY 171 -----ArgLeuPheAsnAsp-----Gly 176  
Db 29321 AAAGATAATCTACATGGTGATCAAAAACCTGCTCGTGATCAACAACACAGCAGTACAACACT 29380  
QY 177 ValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThr 196  
Db 29381 GTAATGATGATGCCAAACTTAAT-----CATGCACACAACAACAGCATTAACT 29428  
QY 197 SerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeu 216  
Db 29429 GATGCTATAATGCAGCGCTCAAGACAGAGGTTGCAACAACATGTTCAACAACCTGCTACT 29488  
QY 217 GluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGln 236  
Db 29489 GAA----- 29491  
QY 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg 256  
Db 29492 -----CTTGATCAGCGGATGGAACATTGAAA 29518  
QY 257 AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAsp 276  
Db 29519 AATAAAGTTGATCAAGTGAATACA-----GATAAGCTCAACCAATTAACACT--- 29566  
QY 277 TyrArgAlaValGluValArgGlyGlyAlaAspLysAlaPheThrThrValAla 296  
Db 29567 -----GAAGCGTCAACTGAT----- 29581  
QY 297 AspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLys 316  
Db 29582 -----AAAAA 29587  
QY 317 AsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgThrLeuAspArg 336  
Db 29588 GAAGCAGTATGATCAAGC-----TTACAAGCT 29614  
QY 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356  
Db 29615 GCAGAAAGCATTTACAGATCCAACTAAT-----GGT 29644  
QY 357 ThrGlnTyrArgPheAspGluVal-----ValPhePheThrIleAspProLysThrAsn 374

[illegible]

```

US-09-815-242-4780
: Sequence 4780, Application US/09815242
: Patent No. US20020061369A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4780
: LENGTH: 7302
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
US-09-815-242-4780

```

Alignment Scores:		0.0138	Length:	7302	
Pred. No.:	181	Matches:	181		
Score:	131.00	Conservative:	157		
Percent Similarity:	32.94%	Mismatches:	346		
Best Local Similarity:	17.64%	Indels:	342		
Query Match:	2.77%	Gaps:	46		
DB:	10				
US-09-914-168-2 (1-919) x US-09-815-242-4780 (1-7302)					
Qy	29	LeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAla	48		
Db	3526	ATTAGTGCACGAATAAATACCTGAAATGAATCCTGATACATAATTAACCAAAAGCCAGCCAA	3585		
Qy	49	IleAsnGlnAlaIleAlaGlyAsnProValLeuLeuThrProGluGlnIleGlnAla	68		
Db	3586	GTGAACAGTGGAGTCTGCA-----TTGAACGGTGTATGAA-----	3621		
Qy	69	ArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsn	88		
Db	3622	AAATAGCAGCAGCAAAACAACTCGAAATCAGATATCGGTGCTTGCACAGACTTGAAC	3681		
Qy	89	PheAspAspGlnSerProIleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAsp	108		
Db	3682	AATGCACACGAACTGCGGCAAACTGCTGAAGTGGATCAAGCACCACAAATCTTGACAGCTGTC	3741		
Qy	109	MetSerValIleGluGluThrThrProLeuSerLeu-----GluGlu	122		
Db	3742	ACAGCGCTAAAAATAAAGCAACATCGTTAAACACAGCGATGGGTAATTTGAAACATGCA	3801		
Qy	123	LeuPheAlaGlnGlnSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyr	142		
Db	3802	CTTGCTGAAAGGATAAATACGAACCTAGTGTCAATTACACAGAT-----	3846		
Qy	143	GlnGlyGlnGlnProAsnSerGluValValValProProThrLeuGluProGluLysPro	162		
Db	3847	---GCGGATCAACCAAAACAAAGCG-----	3870		
Qy	163	GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArg	182		
Db	3871	-----TATGATACCTCGGTTACACAAAGCAGAGCA	3900		
Qy	183	LeuLysAlaLysPheTyrGlnSerGlnSerGlnSerGlyGluThrSerAlaIleGlySerSer	202		
Db	3901	ATT-----ACTAATGCTAATGCGCAAC	3924		
Qy	203	HisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlu	222		
Db	3925	GCGAATGAAACAA-----GTTCAAGCAGCAGCTAACCAATTTGAATCAA---	3969		
Qy	223	SerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAla	242		
Db	3970	GCTAAATACACTGATGATGATGAT---ATAAAGTTGCTCAAGCAAAAGAGTCAGCGAAA	4026		
Qy	243	ArgAlaValGlyTyrTyr-----	249		
Db	4027	CGTGCATTAGTTCATATAGTAACTTGAATATGCGCAATCACTCGAGCAACTAGTCAA	4086		
Qy	250	IleAsp-----LeuSerIleIleArgAsnSerIleGlyGluVal	262		
Db	4087	ATTGACAAATGCAACGACAGTAGAGCGGTAACTGCTGCACAAATACTGCTAATGAATTA	4146		
Qy	263	AspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal	282		
Db	4147	AATACAGCAATGGGTCAACTT-----	4167		
Qy	283	ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeu	302		
Db	4168	---CAAAATGGTATAAATGACCA-----AACACTGTTAAACAAAGTGAACCTTTACA	4218		
Qy	303	IleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAla	322		

```

Qy 603 PheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGly----- 616
      ::::::::::: |||::: |||
Db 5299 CTTGGAATAGTATTGGCAATTAAGATGATGTAAGAGGAGTCAACCATATGTTGATGCA 5358
Qy 616 ----- 616
Db 5359 GATGAGATTAACAATAATGCATACATACAGCAGTTACAAATGCTGAATATATCATTAAT 5418
Qy 617 ----- 617
      ::::::::::: |||
Db 5419 GCAACGAGTCAGCCGACATGATTCATGTCAGTACAGCAAGAGCTAATACAGTGAAGC 5478
Qy 623 GluAspLys-----LysGlyValLysLeuTyrAlaThrLys----- 634
      ::::::::::: ||| :::: ||| |||
Db 5479 ACTACACAAATAGCGGCTTAATAGTGGCACAATACTTACGCAATTAACCAAGAACGACT 5538
Qy 635 -----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
      ::::::::::: ||| ||| |||
Db 5539 GCTAACATCAACCAATTAAGTCATTTAATAATGCTCA----- 5577
Qy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr----- 666
      ::::::::::: ||| ||| ::::::::::: |||
Db 5578 AAGCAAGATTTTAATACGCAAGTACAGATCCACCAATATTAGCAGTAATCAAGTGC 5637
Qy 667 ArgThrLeuGlnHisGluIleSerArgSerIle-----IleGlnAsnGlyGlyTyrAsn 684
      ::::::::::: ||| ::::::::::: |||
Db 5638 AAACTAAAGCTGAACCAATTTGATCAGCAAGCTTTAATACGCAATCCAGAAC 5697
Qy 685 ArgThrTyrSerLeuArgTyrArgLeuAspLysLeuTyrThrGlnAlaProProGluThr 704
      ::::::::::: ||| ::::::::::: |||
Db 5698 AAA-----GATCAAGTGAACAAGT----- 5718
Qy 705 TrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAla 724
      ::::::::::: ||| ::::::::::: |||
Db 5719 -----GTTAACTTTACAGATGCAGATCCAGAAACCAACAGCATACAC 5763
Qy 725 GlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
      ::::::::::: ||| ||| ::::::::::: |||
Db 5764 AATGCGGTA-----ACGGCGCTGAATAATATTATTATCAACGAATGCT----- 5808
Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
      ::::::::::: ||| ::::::::::: |||
Db 5809 -----ACAAATGCCAGAACCAATCAACAA 5829
Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPhe-----GlyAspAsnAla 779
      ::::::::::: ||| ::::::::::: |||
Db 5830 GTTGAACGACGACTTTCACTGTACACACTACGAAACAGCGCTTGATGTGATGAATAA 5889
Qy 780 Tyr-----GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIle 797
      ::::::::::: ||| ||| ::::::::::: |||
Db 5890 GTACACAGATGCTAAACCAATGCAACCAACCAATTAATCTAGCTTA----- 5934
Qy 798 TrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGln 817
      ::::::::::: ||| ||| ::::::::::: |||
Db 5935 -----GATACTTAACCAAT----- 5949
Qy 818 SerLeuArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
      ::::::::::: ||| ::::::::::: |||
Db 5950 -----GCACAAAAAAGTCTCTTACT 5970
Qy 838 Gly-----GlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGlyPheMetLys 855
      ::::::::::: ||| ||| ::::::::::: |||
Db 5971 GGAACACATCATCAAGCCGACAGCTAGCTAGCAAGCAAGCATTCACAAACCCCTGAG 6030
Qy 856 AspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsn 875
      ::::::::::: ||| ::::::::::: |||
Db 6031 GAACCTGAATACAGCGATG-----GTAACCTTGAATAATAGCTTGAAT-----GATAAA 6078
Qy 876 AspThrLysIleGlyAla 881
      ::::::::::: ||| ::::::::::: |||
Db 6079 GACACTACACTTGGCAGT 6096

```

```

RESULT 24
US-09-815-242-8898
; Sequence 8898, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: EITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8898
; LENGTH: 18846
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(18846)
US-09-815-242-8898

Alignment Scores:
Pred. No.: 0.0552 Length: 18846
Score: 131.00 Matches: 181
Percent Similarity: 32.94% Conservative: 157
Best Local Similarity: 17.64% Mismatches: 346
Query Match: 2.77% Indels: 342
DB: 10 Gaps: 46

US-09-914-168-2 (1-919) x US-09-815-242-8898 (1-18846)
Qy 29 LeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAla 48
      ::::::::::: ||| ||| ::::::::::: |||
Db 5863 ATTAGTCGACAGCAATTAACCTGAATGAATCTGATACCAATTAACCAAAAGCGCCAA 5922
Qy 49 IleAsnGlnAlaValAlaGlyAsnProProValIleLeuThrProGlnGlnIleGlnAla 68
      ::::::::::: ||| ||| ::::::::::: |||
Db 5923 GTGAACAGTGCAGACTGTGCA-----TTGAACGGTGATGCA----- 5958
Qy 69 ArgLeuAsnAlaValGlyLeuAsnAlaLysProGlnIleSerGlnAlaLeuAspValAlaAsn 88
      ::::::::::: ||| ||| ::::::::::: |||
Db 5959 AATTTAGCAGCGACGCAAAACCAACTCGAANAATGAGATATCGGTCGTTTGACAGACTTGAAAC 6018
Qy 89 PheAspAspGlnSerProIleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAsp 108
      ::::::::::: ||| ::::::::::: |||
Db 6019 AATGCACAACAGCAATGCGGCAAAATGCTGAAGTGAATCAAGCAACCAATCTGCAGCTGC 6078
Qy 109 MetSerValIleGluGluThrThrProLeuSerLeu-----GluGlu 122
      ::::::::::: ||| ||| ::::::::::: |||
Db 6079 ACAGCGCTAAATAAATGAAGCAACATCGTTAAACAGAGCGATGGTAATTGAAACATGCA 6138

```

Oy	123	LhePheAlcGlnGlnSerSerGlnMetGlyIleAsnProAsnAspArgTleProGluTyr	142
Db	6139	CTTGCTGAAAGGATATATACGAACCGTAGTGCAATTTCACAGAT	6188
Oy	143	GlnGlyClnProAsnSerGlnValValAlaProProThrLeuGluProGluLysPro	1626
Db	6184	---GGCGATCAACCAAAACACACACG---	620
Oy	163	GlyLeuIleLeuArgLeuTyrLalaArgLeuPheAsnAspGlyValAsnLysValProArg	1822
Db	6208	-----ATGCTACTGGCGTTTACACAGACAGAAACA	6233
Oy	183	LeuValAlaLysPheTyrGlnSerGlnSerGlnSerGluTyrThrSerAlaIleGlyIleSer	202
Db	6238	ATT-----ACTAATGCTAATGGCAGCAAC	626
Oy	203	HisGlnLysThrGluProTyrAlaAsnIleLeuAlaIleGluGlnAspIleThrGlnGlu	222
Db	6262	GCGAATGAAACACAA-----GTTCCACACGACTAAACCAATGGATCAAA---	630
Oy	223	SerAlaMetAspLeuAsnGlySerTleProArgLeuArgGlnThrAlaLeuValAlaAla	242
Db	6307	GCTAAATAATGACTTGAATGGTGAT---AATAAAGTTGCTCAAGCAAAAGAGTCAAGGAAA	636
Oy	243	ArgAlaValAlcTyrTyr-----Asp	249
Db	6364	CGTGATTAAGCTTCATATAGTAACCTTGATATGCGCATCAACTCGACAACTAGTCAA	642
Oy	250	IleAsp-----LeuSerIleIleArgAsnSerIleGlyGluVal	262
Db	6424	ATTGACAAATGCAACAGACAGTACAGCGGTAACTGCTGCACAAATATCGCTAATGAATTA	648
Oy	263	AspValIleIleHisAspLeuGlyLysProValTyrIleAspTyrArgAlaValAlcVal	282
Db	6484	AATACAGCAATCGGTCAACTT-----	650
Oy	283	ArgGlyGlnGlyAlaAspAspLysAlaIlePheThrThrValAlaAspGluValProLeuLeu	302
Db	6505	---CAAAATGCTATTAATATGACAAA-----AACCTGTTTAAACACAAAGCAATTTC	655
Oy	303	IleGlyAspValPheHisHisGlyTyrGlnThrLysLysAsnLeuIleGluAsnAla	322
Db	6556	GATGCTGAC-----CAAGGTAAAGAAAGATGCTTACACAAATGCTGTATACAAATGCT	660
Oy	322	SerAla-----GlnHisGlyTyrPheAspArgTyrGThrLeuAspArgSerVal	338
Db	6607	CAAGGTATTTAGATTAAGCACACCGGTCAA-----AATATGACGAAGACACAAAGTT	665
Oy	339	AspValIleLeuProAsp---AsnThrAlaAspValSerLeuIleTyrAspThrGlyThr	357
Db	6658	GAAGCTGCATTAATTAACAAGTAACGACGCTCAACAAATGCTTTAAACGGTACCAAAATGTA	671
Oy	358	GlnTyrArgPheAspGlyValValPhePheThrIleAspProLysThrAsnIleuThr	377
Db	6718	AGACAAACA-----AAATCAGATGCGAAAGCAAACTTAGTGACA	675
Oy	378	ThrAspProAspLysLeuProValLysArgGlyLeuLeuGlnGluLeu-----Leu	394
Db	6757	TTAACACACCTTAATATATGCACAAAACAGATTTAATCATCAAAATTCGAAGCGCAACA	681
Oy	395	ThrValAsn-----MetGlyGluAlaTyr	402
Db	6817	ACACTCAACCGCTGAATGCTGTTAAACGAAACACAGCAAGCACTTAGATGTGTCATGCA	687
Oy	403	AsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMet	422
Db	6877	CGATTCAATCACTAGCTATAGCAAAATAAAGATCAAACTAAAGCGACGAA-----AATTAC	693
Oy	423	ValAsnThrGlnIleValPheProGluTyrGlnGlnIleGlnAsnAspGlnValSerPhe	442
Db	6931	ATCACACGACAT-----CCAATAAGAAACAGCACTTTGATTAATGCTATATCAACA	698
Oy	443	GlnClnSerSerSerSerArgThrGlnProAlaGlnValAspGluSerThrLeuClnPro	462

[illegible]





Dh 796 AGCCAAACAATTCAACAATATTAAGCTAAACGTCATCACTTGATTAATGATGTTGAAATCG 855  
Qy 302 -----LeuilegIyAspValPheHISISGLySTyTgLUthrlYsAsn 317  
Dh 856 CTTAATATTCATTAATTAAGCTAAAGCTGACGACCTTTGAAAAATATACGGACATTAAT 915  
Qy 318 LeuilegIyAsnAlaSerAlaGluHISGLyTyrPheAspGlyArgTyrLeuAspArgSer 337  
Dh 916 GTTTAGCAAGACGTAAACAAAATCAATCTGAACCAAAATGACGATATGAGAACAA 975  
Qy 338 ValAspValIle-----LeuProAspAsnThrAlaAspValSerLeuileTyrAspThr 355  
Dh 976 GAAATTTAATTAATGACCTTTTGAATAATATATCAATGAGATTTCGAAAGCTCAACACT 1035  
Qy 356 GlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGln 375  
Dh 1036 -----TATAAGTCTCGAAAGAAAGTAAACAAAACAA 1065  
Qy 376 LeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeuThr 395  
Dh 1066 CTCACCGCT-----GTCATTCGTCAATTAAGAACCACTATATGTT 1107  
Qy 396 ValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIle 415  
Dh 1108 TCACACGACAGCACACGATGAAAAATTGACAAATTA----- 1146  
Qy 416 AlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIle 435  
Dh 1147 ---AACCAATCACTATACATTAATGTCA-----GACCAATCAAGATGTT 1185  
Qy 436 GlnAsnAsp-----GlnValSerPheGluGlnSerSerSerSerArgThrGlu 451  
Dh 1186 AACCATATATTCCTTTTAAACCATATCTATAGAAAGAAATGAGGCTAAAAAATCAAGA 1245  
Qy 452 ProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471  
Dh 1246 CTA-----GATTCTCGATTAGTTGAAGTCTTTGACCAATTCGAAGATATTCAG 1293  
Qy 472 GlyIleLeu-----MetAspIleSerProIle 480  
Dh 1294 GGTCAAAATTAACACGACTAAAAAAGAAATATCAACAGCAACCAAAAGAACTTCTGCTGA 1353  
Qy 481 GlnPheSerAlaSerAsnLeuileGlnAspLysLeuAsnLeuValAlaAlaValAlaArg 500  
Dh 1354 GATTAAGAAATTAATAAATATATGAAGAAAGCCTCTACTATCCAAAAAAGCAAAACGAA 1413  
Qy 501 His-----LeuTyr-----AspMetProAspAspArgVal--- 510  
Dh 1414 TACGAACAGCAAAATGTATCAAGCATATCGATATACGAAAAAATGAACACGATATGAT 1473  
Qy 511 ---LeuAlaIleAsnHISAspAsp-----GlyVal----- 519  
Dh 1474 AGTTTGGCAACCCAGAGGAAGATATATATTTTTCATATGGCTGCAACATATTTTG 1533  
Qy 520 -----AsnArgSerIleLeuGly-----ArgIleSerAspAla 530  
Dh 1534 AAAGCTAAATTAAGATTAAGAGATATTCATGCTGACGTTGCGGAATATATGATGATG 1593  
Qy 531 ValSerAlaValAlaArgAlaIle-----LeuProAspGluSerGlnGluVal 547  
Dh 1594 CCATCTAAATTAACCTGAGCAATTTGAACAGCATTTAGTCTTCATTAACAACATGTTATT 1653  
Qy 548 IleAsp-----LeuProGluArgThrAlaLeu 556  
Dh 1654 GTAGATTTCACAAAAGATGAGCCACCGATTCATTAATTTTAAAGAACGTAATTTAGGT 1713  
Qy 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeu----- 573  
Dh 1714 CGTGCAGACCTTTTACCATTAATGTTATATACAGATGAGTGTAGCGACTGATATTAAAG 1773  
Qy 574 -----TyrValPheValAlaSerAspLysProArg----- 583  
Dh 1774 TCTATTGCTTAAAGAGCAACGATTTATATGATTGCTTCGGAAGACGTTAAAGTACGA 1833

Qy 584 -----AspGlyGlnIleGly----- 588  
Dh 1834 CCAGATATTCAAAATATTTATTCGGAAATTTATAGTAATACGATTAATGCTGATTTA 1893  
Qy 589 -----LeuGlyTyrPheLysAspThrGlnTyrArgLeuValThrLysPhe 603  
Dh 1894 AAGCATGCAAAATGAAATTCGACGTCGATTAATATTCGAACCTGATATGTTACT---TTG 1950  
Qy 604 GlnHISAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlu 623  
Dh 1951 GAAGGTGATATGTAACCTCGGTGTCTATGACGTGGTGGCGCTGTAAAGTCAAAA 2010  
Qy 624 AspLysLysGlyValIleLeuTyrAlaThrLysProLeuSerHISProLeuAsnAspGln 643  
Dh 2011 AGTATTTGCTGTCAAAAAGACGAGTTCACA---ACAAAGACACCACTTTGAAGATTAAC 2067  
Qy 644 LeuArgAlaThrLeuGlyTyrGlnGlnGluValAlaPheGlnHISSerThrAsnGlyPheAsp 663  
Dh 2068 TTGCCTCAACAGAAATCATTTGAAACACAACTTAAGAAGTTGAAGATTAAGATGATCA 2127  
Qy 664 LeuSerThrArgThrLeuGluHISGluIleSerArgSerIleIleGlnAsnGlyGlyTyr 683  
Dh 2128 TTAAGTGAAGCTGATTTTGAAGAAAGTCAAAAGCATATATACCTTAA----- 2175  
Qy 684 AsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699  
Dh 2176 GAGCAAGTCATCAATTTTGAAGATGAGCTCGATAGATTAACTACACAA 2223

RESULT 26  
US-09-815-242-8027  
Sequence 8027, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlser, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OR INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011a  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8027  
LENGTH: 3567  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(3567)  
US-09-815-242-8027  
Alignment Scores:

Pred. No.: 0.00542 Length: 3567  
 Score: 130.50 Matches: 143  
 Percent Similarity: 35.31% Conservative: 131  
 Best Local Similarity: 18.43% Mismatches: 287  
 Query Match: 2.76% Indels: 215  
 DB: 10 Gaps: 35

US-09-914-168-2 (1-919) x US-09-815-242-8027 (1-3567)

QY 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValAlaAsnPheAspSerGlnSer 93  
 DB 94 GGTCAAAATGCAAGCGGTAAAGTATATTACAGATGCTTTAAATGG----- 141  
 QY 94 ProILeserArgIleGluGlnSerProProLeuGluLeuAspMetSerValIleGlu 113  
 DB 142 -----GCTTGGCGCAACATCGCTAA-----TCATTAGCTGCTCAAAAATGGAA 189  
 QY 114 GluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGlnSerThrGluMetGlyIle 133  
 DB 190 GATATTATCTTCTCAGTGCAGAAACATCGCAAGCTCAAAATTATGCTGAAGTACAGTTA 249  
 QY 134 AsnProAsnAspTyrIleProGluTyrGlnGluGlnProAsnSerGluValVal 153  
 DB 250 AGATTGATATATCATCTTAAAGCTCAGTGTGATGAA-----AACGAAGTTATTGTA 303  
 QY 154 ProProThrLeu-----GluProGlu----- 160  
 DB 304 ACAAGAAGATTGTATCGAAGTGTGAAGAGTACATACATATATATGACCGTCAGAGA 363  
 QY 161 -----LysProGlyLeuIleLysArgLeuTyrAlaArg 171  
 DB 364 TTTAAAGATATTCGCGATTTATTTTAGATTCTGATTGGGAAAGACATATATAGCAT 423  
 QY 172 LeuPheAsnAspGlyValAlaAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSer 191  
 DB 424 ATCTCCCAAGTGTAGAGTTGTGAATA-----CTAATGCTAAACCAATTGATAGACGT 477  
 QY 192 GlnSerGlyLysThrSerAla-----IleGlySerSerHisGlnLysThrGluProTyr 209  
 DB 478 CAATATTATTCAGAAATCGCTGCTGACTTAATATTAATAAGCTAAAGCTAATCATTA 537  
 QY 210 AlaAsnIleLysAlaAlaLeuGluAspIleThr-----GlnGluSerAlaMetAspLeu 227  
 DB 538 AATTAACCTTGACCAACAGAGATTAATTTAACGAGCTAACGACATTTATATGATTTG 597  
 QY 228 AsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr 247  
 DB 598 GAAGTTCGCTAGAACCTCTAAAGAGGAGCGACCTATAGCTAAA----- 642  
 QY 248 TyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle----- 265  
 DB 643 -----GAATATAAGACACTTTCGCACCAATGAACATAGTACATTGTGTGTACA 693  
 QY 266 IleHisAspLeuGlyLysProVal-----TyrIleAspTyrArgAlaValGlu 281  
 DB 694 GTTCACAGATATTATCATATTTACAAATGACAAATAGACAAATTTAGCTTTAAATGAT 753  
 QY 282 ValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu 301  
 DB 754 TTTCAGAGGTCAACAGCAATTAAGAA-----GCTGACAGCAACAGTTTA 798  
 QY 301 ----- 301  
 DB 799 AGCCAACAATTCACAAATTAAGTAAAGTCAATCATCTGATATGATGTTGATCG 858  
 QY 302 -----LeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsn 317  
 DB 859 CTTAATATTATCAATTAAGTAAAGCTACGGAAGCCTTTGAAAAATTTAGCGGCAATTAAT 918  
 QY 318 LeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgThrLeuAspArgSer 337  
 DB 919 GTTTTGAAGACAGTAAAGAAATTCATCTGAAACAAATGCGCATATATGAAGAACAA 978

QY 338 ValAspValIle-----LeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355  
 DB 979 GAAATTTTAATTGAGCTTTTAGAAAATATATATCAAAATGAGATTCTGAGCTCAAGTACT 1038  
 QY 356 GlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGln 375  
 DB 1039 -----TATTAAGTCTCTGAAAGTAAACAAAGAA 1068  
 QY 376 LeuThrThrAspProAspLysLeuProValLysArgGluLeuGlnGlnLeuThr 395  
 DB 1069 CTCACAGCT-----GTCATTGCTGATTAAGAAACCAACTATATGCTT 1110  
 QY 396 ValAsnMetGlyAlaAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIle 415  
 DB 1111 TCACAGCAGACACGACGATGAAAATTTGCAACAATTAA----- 1149  
 QY 416 AlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIle 435  
 DB 1150 ---AACCAATACATATATATATGCTCA-----GAGCATTCAGATGCTT 1188  
 QY 436 GlnAsnAsp-----GlnValSerPheGluGlnGlnSerSerSerArgThrGlu 451  
 DB 1189 AACATGATATTCGTTTTTAACCATCTATTAAGAAAGTGAAGCTAAAAAATCAAGA 1248  
 QY 452 ProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471  
 DB 1249 CTA-----GATTCGATTTAGTTGAAGCTTTTGAACAATTGAAGATATTCAG 1296  
 QY 472 GlyIleLeu-----MetAspIleSerProIle 480  
 DB 1297 GGTCAAAATAAAGACACTAAAGAAATATCAACAGACCAACAAAGAACTTTCGCTGTA 1356  
 QY 481 GluPheSerAlaSerAsnLeuIleGlnLysLysLeuAsnLeuValAlaAlaLysAlaArg 500  
 DB 1357 GATTAAGAAATTAATAATATATAGAAAAGACCTCACTGATATCAAAAAAGCAAAAGCAA 1416  
 QY 501 His-----LeuTyr-----AspMetProAspAspArgVal-- 510  
 DB 1417 TACGAGAGAAATTTGATATCAGCATATATGATATCCGAAATAAGAAACGATATTGAT 1476  
 QY 511 ---LeuAlaIleAsnHisAspAsp-----GlyVal----- 519  
 DB 1477 AGTTTGCACAGCAAGAGGAAGATATATCTATTTTTCAAATGCGCTCAACATATTTTG 1536  
 QY 520 -----AsnArgSerIleLeuGly-----ArgIleSerAspAla 530  
 DB 1537 AAACCTAAATTAATAAGATTTAAAGGATATTCATGCTGACGTTGGGGAATTTATGATGTG 1596  
 QY 531 ValSerAlaValAlaArgAlaIle-----LeuProAspGluSerGluAsnGluVal 547  
 DB 1597 CCATCTAAATTAATCACTGAGCAATTTGAACACAGCATTAGCTGCTCATTTACACATGTTAT 1656  
 QY 548 IleAsp-----LeuProGluArgThrAlaLeu 556  
 DB 1657 GTAGATTTCAGAAAAAGATGACGCCAGCTATTCATTTTAAAGAAACGTAATTTAGCT 1716  
 QY 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeu----- 573  
 DB 1717 CGTCGACGTTTTTACCATTAATGTTTATACAGAGTAGAGTGGTACGACATATTAAAG 1776  
 QY 574 -----TyrValPheValAlaSerAspLysProArg----- 583  
 DB 1777 TCTATTGCTAAAGAGCAACGAGATTTATAGTATGCTTCGGAAACGTTAAAGTACGA 1836  
 QY 584 -----AspGlyGlnIleGly----- 588  
 DB 1837 CCAGAAATATCAAAATATATTCGGAATTTATAGTAATATACAGATTATTCGTTGATCATTTA 1896  
 QY 589 -----LeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPhe 603  
 DB 1897 AACGATGCAATGATATTTGGCAGTCGCTTAAATATGCAACCTGATTTGTTACT---TTG 1953  
 QY 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlu 623

Db 1954 GAAGGTGATATTGTAATATCTGCTGTTTATGACTGGTGGCGCTGTAAGTCAAA 2013  
OY 624 AsplyslysglyvallyleuylalathrlsProleuSerHisProleuasnpsgl 643  
Db 2014 ACTATTCTCTCTCAAAAGACGAGTTGACA--ACAATGACACCATTTAGAGATTAC 2070  
OY 644 LeuATGAlathrlleuGLYTrGInGlnuvalrhegIyHisSerThrAsnGlyPheasp 663  
Db 2071 TTCGCTCAACGATCATCTTTGACACACAAATTTAAAGAGTTGAAGTAAAGATCA 2130  
OY 664 LeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrp 683  
Db 2131 TTAAGTGAAGCTGATTTTGAAGAAAGATCAAAAGCATATACCTTA-- 2178  
OY 684 AsnArgThrTySerLeuArgTrArgleuAspIlyleuLysTrGln 699  
Db 2179 GAGCAAGTCATCATTTTGAATGAGCTCGATGATTAACTACACAA 2226  
RESULT 27  
US-09-815-242-4779  
Sequence 4779, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 4779  
LENGTH: 7311  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
US-09-815-242-4779  
Alignment Scores:  
Pred. No.: 0.0155 Length: 7311  
Score: 130.50 Matches: 164  
Percent Similarity: 32.91% Conservative: 123  
Best Local Similarity: 18.81% Mismatches: 314  
Query Match: 2.76% Indels: 271  
Db: 10 Caps: 44  
US-09-914-168-2 (1-919) x US-09-815-242-4779 (1-7311)  
OY 26 SerGlnAlaLeu-----AlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisVal 42  
Db 271 TCACAGCTATTATGATGTAAGTCAAAACATTAACGAGCGTAAATTACTTAATGSCAGAT 330

OY 43 ProAlaHisAspThrAlaIleAsnGlnAla----- 52  
Db 331 CCTACACAAACAAATGATATTAACAGCAGTAGACAAAGCTGACGATTTATGATTA 390  
OY 53 LysAlaGlyAsnProProValIleuLeuThrProGlnGlnIleGlnAlaArgleuAsnAla 72  
Db 391 CAAAGTGCTACTATGATGATCAACAGCAAGTGAAGCATCTAATGAGTGAAGCAGA 450  
OY 73 AlaGly-----LeuAsnAlaLysProGlnSer 81  
Db 451 GCGAAACACCATTAATGTTAATGACAAATTTGCAAAATGCAAAACACAGCAAAACAA 510  
OY 82 GlnAlaLeuAspValValAsnPheAspAsp----- 91  
Db 511 CAATTGGCGCACTTAACACACTTAATGATGCGCAAAACATCATTTGAAGTCAAAAT 570  
OY 92 ---GlnSerPro-----IleSerArgIleGlyGlnGlnSerProProleuGly 106  
Db 571 ACAGAAAGCGCCACTTGTACAGATGTCACGATTAATCAAAAGACAAACGTTAGT 630  
OY 107 -----LeuAspMetSerValIleGlnGlnThrThrProleuSerleuGln 121  
Db 631 CATGCGATGATTAATTAGAAATAGCTGTCGCGATATCAAAAGCAATTAACGCTGGA 690  
OY 122 GluLeuPheAlaGlnGlnSerThrGlnMetGlyIleAsnProAsnAspIlyIleProGlu 141  
Db 691 GATTAT-----CATGATGCAACATGCGCA-----AGCAAAATGACTTAACCAAGCT 738  
OY 142 TyrgInGlyGlnGlnProAsnSerGluValValProProThrleuGlnProGluLys 161  
Db 739 GTAACAGCTGCTAATAATATCATTTAATCAAACTACATGCGCTACGATTAATCA----- 792  
OY 162 ProGlyLeuIleLysArgleuThrAlaArgleuPheAsnAspGlyValAsnLysValPro 181  
Db 793 -----CATGATGTTAAT----- 804  
OY 182 ArgleuLysAlaLysPheTyrgInSerSerGlnSerGlyGlnThrSerAlaIleGlySer 201  
Db 805 -----GTTGCAAGCAACAAAGTAAATAT 828  
OY 202 SerHisGlnLysThrGluProTyrgAlaAsnIleLysAlaIleGlu----- 217  
Db 829 ACGAAAGTTGCATTAGATGATGATGAAACCTTCACAGCAGCTAAACAAAGCAACAC 888  
OY 218 ----- 219  
Db 889 AGACTGATCATTAATGATCATTTGATTAATGCGCAAAAGCAAGTCAATCAACAAT 948  
OY 220 ThrGlnGlnSerAlaMet---AspLeuAsnGlySerIleProArgleuArgGlnThrAla 238  
Db 949 ACGCAATCATCTGATTAATGCTGCGAGTTAATGTCAC-----AAACAAACAGCA 996  
OY 239 LeuValAlaAlaGlyAlaValGlyTyrgTrAspIleAspLeuSerIleIleArgAsnSer 258  
Db 997 GAACTTTAAATACCTGCGATGGGT-----AACTTA 1026  
OY 259 IleGlyLysValAspValIleIleHisAspLeuGlyGluProValTyrgIleAspTyrg 278  
Db 1027 ATTAATGCGATT-----GCAATCATCA 1050  
OY 279 AlaValGluValArgGlyGlu-----GlyAlaAsp---AspLys-----AlaPheThr 293  
Db 1051 GCGGTGAAACAGCTGTAATCTCATCAATGATGATGATGATGATGATGATGATGAT 1110  
OY 294 ThrValAlaAspGluValProleuLeuIleGlyAspValPheHisIleGlyLysTyrgLys 313  
Db 1111 ACAGCGTAAATGATGACACAGCAAGATGATTAACAAACAAAGTGCATAAATGCGAACCA 1170  
OY 314 ThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrgPheAspIlyArgTrp 333  
Db 1171 ACGAG-----GTAGACAGCTATTACTAA----- 1197  
OY 334 LeuAspArgSerValAspValIleleuProAsnThrAlaAspValSerleuIleTyrg 353

```

Db 1198 GTTCAAAACACACTTCAAGCGTTAAATGAGACCATTAATTACAAAGTTGCT-----AAA 1251
Oy 354 AspThrGlyThrGlnThrArgPheAspGluValValPhePheThrIleAspProLysThr 373
Db 1252 ACAAAATCGCAGCAGCA-----GCAAATGATGCTTTTAA 1284
Oy 374 AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluLeu 393
Db 1285 AGC-----TTAAATGATCCCTCAAAAACACATTTAAAGACCAAGTTACAGCTGCAACT 1338
Oy 394 Leu-----ThrValAsnMetGlyGluAla 401
Db 1339 TTGATTAAGTCAAGTTCATCAAAATGCAAAATGCGAATACGCTTAACCAAGCA----- 1392
Oy 402 TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsn 421
Db 1393 -----ATGATGATGTTTAAAGCAGACGATTCACATTAACGCAACACTAA----- 1437
Oy 422 MetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnVal 440
Db 1438 ---GCAATATGCAAAATATATCAACAGACGATCAACCAACCAAACTATGATCAACT 1494
Oy 441 SerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 1495 GTTCAACCCCGCAATATATTTATCATGCAACAACTGCAACATTAAGATTAATATGCGATT 1554
Oy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMet---AspIleSerPro 479
Db 1555 AATCAACAGCAGCAGCACTGTGATACACAGCAACAGCAGATTCATGCTGATGTC----- 1608
Oy 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAla 499
Db 1609 -----AAGTTACAA-----AATGATTAAGATCATGCT 1617
Oy 500 ArgHisLeuThrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
Db 1618 -----LeuGlyArgIleSerAspAlaValSerAlaValAlaAlaArgAla 537
Oy 520 AsnArgSerIle-----LeuGlyArgIleSerAspAlaValSerAlaValAlaAlaArgAla 537
Db 1636 AAGCAAAACGGTTAGTCACTATTTAGCATCTTAACATGACCAAAAACATATGGAAGATACG 1695
Oy 538 IleLeuProAspGluSerGluAsnGluValIle-----AspLeuProGluArgThrAla 555
Db 1696 TTAATGATGATGTAACAACTAGTCAAGACAGTTAAGCAAGATTGACTGAGCAACCAACGA 1755
Oy 556 Leu-----AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLys 570
Db 1756 TTGATCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1815
Oy 571 ValProLeuThrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGly 590
Db 1816 AGCAGTGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1872
Oy 591 TyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg 610
Db 1873 -----CAAAATGCTGAGTCAATGATGATGATGATGATGATGATGATGATGATGATG 1923
Oy 611 AspGlyTyrGlnAlaGlyAlaGluLeu----- 619
Db 1924 GGTAAATGATCACTGACGACACAGATTAATCATCTAAATAATGATGATGATGATGATGATG 1983
Oy 620 ---ArgLeuSerGluAspLysLysGlyValLysLeuThrAlaThrLysProLeuSerHis 638
Db 1984 GAACGATTAAGTCAAGATTAACCA-----ACTGCTGGAATATCTCTAATATAT 2031
Oy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlyHisSer 658
Db 2032 -----TTAGATCAATTAACACAGCT-----CAACAAACAAAGCCTGGAATAATCAA 2076
Oy 659 ThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleLeu 678

```

```

Db 2077 ATTAAT-----AATGCAACAACTCGTGGCGAAGTAGACCAAAAATTAAC 2121
Oy 679 GlnAsnGlyGlyTyrPheAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuThr 698
Db 2122 GAA-----GCACACGACTTTAAACCAAGCAATGGAAGCTTTAGGTAT 2163
Oy 699 GlnAlaProProGluThrTyrPheGlnAspLeuProValAspPheValAsn---GlyLysPro 717
Db 2164 AGCATTAAGATCAACACAGCAACGACGCGGTAGACAGTTATATCAATGAAGATTAACCA 2223
Oy 718 SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 2224 CAAAAAGATGCTTACCAAGCAGCAGCACTCAAAATGCAAAAGATTTAATTAATCAAACTAAC 2283
Oy 728 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal 757
Db 2284 AATCCAAAGCTTTGATTAAGCAGCA-----GTT 2310
Oy 758 SerAspAlaAsnMetAlaIleAlaAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 2311 GAACATTTGACACAGCAAGCTGTTAACCAGCTTAAGATTAAC-----CTACACGCTGAT 2361
Oy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThr 789
Db 2362 CAAAAACTTGACAGCATTAACAAACATGCGGTACT 2397

RESULT 28
US-09-954-456-1807
Sequence 1807, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
FILE REFERENCE: 689290-76
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/09/954,456
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1807
LENGTH: 3336
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1807

Alignment Scores:
Pred. No.: 0.00781
Score: 128.50
Percent Similarity: 33.76%
Best Local Similarity: 19.89%
Query Match: 2.72%
Indels: 239
Gaps: 46

```



```

Qy 713 ValAsnglyLysProSerGlnGluAa-----LeuLeuAlaGly 725
    |||||
Db 2154 GCCAATGGACCGCATTTTCATGGAAGTTCTGTAAGTTTGACAAATACGATGCCAGAGC 2213
Qy 726 ValAlaValHisLysThrValAlaAlaAspAsnLeuValAspAsnProMetArgGlyTyrGln 745
    |||||
Db 2214 ACAGCGGTGGAAGGACCTTCAGGAATGAAATGACCCTTACCTTACGACCCACTGAA 2273
Qy 746 ArgTyrSerLeuGluValAlaGlySerGlyLeuValSerAspAlaAsnMetAlaIleAla 765
    |||||
Db 2274 TCTGGAGGTGCTGCAGCGCTTTCACAAAGAGGACAGTGTCCAGGCTTCACAGATTCCGA 2333
Qy 766 ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAla 785
    |||||
Db 2334 GCAGCCACTGGAGTGGCATCTGTTCTTCAGGCGAGCTC-----GGAACCATGAGAACA 2387
Qy 786 HisGlnMetThrGlyGlyIleGlnAlaGlyTyr-----IleTyrSerAspAsnIle 802
    |||||
Db 2388 AGGCATTGCCACTGGAGGAACCAATAGAGACTAGCTGATGGGCGCATGATGAATTTT 2447
Qy 803 AsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyr 822
    |||||
Db 2448 CTGCACTCTTACTTTTCAGAAACATTTGCTGCGGAGGAGGACATGCCAGCA 2507
Qy 823 AlaHisAspSerLeuSerProIleSerAspLysGlyTyr--LeuThrGlyGlyGlnVal 841
    |||||
Db 2508 GCAATGACTGCTGCTGTGATCATGATGATGAAGCGGAGATGCCATGCTCTCTCGTG 2567
Qy 842 LeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaVal 861
    |||||
Db 2568 GGCCTCGGCGGTGTTGAGATTATTTGCTGATGACCTGGATGACGACTTCTTGAGCTCA 2627
Qy 862 PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAla 881
    |||||
Db 2628 CTGGA-----CCCAATTTTAAAAAACTTCGACAGATTAAGCCTT 2666
Qy 882 GlyValGlyValArgTyrPalAserProValGlyGlnValArgValAspValAlaThrGly 901
    |||||
Db 2667 GGTGTGTGATGTGA-----GCCAAGAGATTCAGCACCTTAAGACACGCGTTATGGG 2723
Qy 902 ValLysGluGluGlyAsnProIleLysLeu 911
Db 2724 ATTGAAATCCTGTGCCATTCATAGAAGTC 2753

RESULT 29
US-09-881-752A-307
: Sequence 307, Application US/09881752A
: Patent No. US20020115078A1
: GENERAL INFORMATION:
: APPLICANT: Kleantous, Harold
: APPLICANT: Al-Garawi, Amal
: APPLICANT: Miller, Charles
: APPLICANT: Tomb, Jean-Francois
: APPLICANT: Oomen, Raymond P.
: TITLE OF INVENTION: Identification of Polynucleotides
: TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
: FILE REFERENCE: 06132/041002
: CURRENT APPLICATION NUMBER: US/09/881,752A
: PRIOR FILING DATE: 1997-04-01
: PRIOR APPLICATION NUMBER: US 08/833,457
: NUMBER OF SEQ ID NOS: 370
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 307
: LENGTH: 5832
: TYPE: DNA
: ORGANISM: Helicobacter pylori
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (387)...(5777)
: US-09-881-752A-307

```

```

Alignment Scores:
Pred. No.: 0.0223
Score: 127.50
Percent Similarity: 33.90%
Best Local Similarity: 19.24%
Query Match: 2.70%
DB: 10 Gaps: 52

US-09-914-168-2 (1-919) x US-09-881-752A-307 (1-5832)
Qy 55 GlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaAlaGly 74
    |||||
Db 2130 GGAGCGCCAGAAATCATTTGCGCAATGAAGAA----- 2162
Qy 75 LeuAsnAlaLysProGlnSerGlnAlaLeuAspValAlaAsnPheAspAsnPro 94
    |||||
Db 2163 -----AGGATGTTGCAAGAACCGGTGATGCTTTTGATATACGCGCACACT 2213
Qy 95 IleSerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGluGlu 114
    |||||
Db 2214 AATGGCGTTAAAGGGGCTAACAAAGCCCTTTA----- 2246
Qy 115 ThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsn 134
    |||||
Db 2247 -----AAATCGCTCAGTGAATCATTAAGGCAACAGGCGGCTTTCAGGCAAAAC 2297
Qy 135 ProAsnAspTyrIleProGluTyrGlnGlyGlnProAsnSerGluValAlaValPro 154
    |||||
Db 2298 CTTTAAAGTAACCGCGTGTATTTTCAGGC-----AGAAGCGTGAATGGTGGG 2348
Qy 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174
    |||||
Db 2349 CCTAATCTCAAAATGATGATGATGCGGCTTGCCATAAAACATGCGCTTAGACCTTCACA 2408
Qy 175 AspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly 194
    |||||
Db 2409 -----CCGATTTGTTATCCAACTT-----GAGAGAGAGGC 2441
Qy 195 GluThrSerAlaIleGlySerHisGlnLysThrGluProTyrAlaAsn--IleLys 213
    |||||
Db 2442 TATGCCACCAAGCTCAACAGCGCTAAACCATGATGATGACCAAAAGCAATGACATGAG 2501
Qy 214 AlaAlaLeuGluAspIleThrGlnGlnLysAlaMetAspLeuAsnLysSerIleProArg 233
    |||||
Db 2502 GAGTGCTTCAGAAATCATCAGAGGGGTATCCGGTCTACTCAACCGCGCT--CCTAAC 2558
Qy 234 LeuArgGlnThrAla-----LeuValAlaAlaArgAlaValGly 246
    |||||
Db 2559 TTGCACACAGCAATCCATTCACAGCTTCATCCAAAGCTGATTCGCGCAAGGATCCAA 2618
Qy 247 TyrTyr-----AspIleAsp----- 251
Db 2619 TTGCACCGCTTAGTGTTGTTACAGCGTTCAACGCCGATTTTGACGGGAGCAAAATGGCGGTG 2678
Qy 252 --LeuSerIleIleArgAsnSerIleGlyGlu-----Val 262
    |||||
Db 2679 CATGTCCTTTAAGCAGAGGATCCCTGAATGACAGGTGCTGATGCTAAGCTGTATG 2738
Qy 263 AspValIleIleHisAspLeuGlyLysProValTyrIleAsp----- 276
    |||||
Db 2739 AATATCCCTTTGCTGCTAGCGGTAAAGCGGATGCTTCAGCCAGACATATGTTTAA 2798
Qy 277 -----TyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThr 293
    |||||
Db 2799 GGGCTTTATATCTTTCTTTAGAAAAGAGCGGGGTCAAGCGGAGCATTAAGCTTTTCT 2858
Qy 294 ThrValAlaAspGluValProLeuLeu--IleGlyAspValPheHisIleGlyTyr 312
    |||||
Db 2859 ACGGTGAATGAATCATCATCCGCAATTTGACAGAAAGATTAAGATCCACCAAGAT 2918
Qy 313 GluThr-----LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAsp 330
    |||||
Db 2919 AGGTTTATAGATCAAGGAATATATTCCTACGAGTGA----- 2957

```

```

OY 331 G1YAGTTPLeuAspArgSerValIleLeuProAsp----- 344
    ||||| :||| :|||
Db 2958 GCGCGATGATCATTAAGTCC-----ATTTGCCGATTTATCCCTACGGATTGG 3008
OY 345 -----AsnThrAlaAspValSerLeuIleTyrAsp----- 354
    ||||| :||| :|||
Db 3009 TCGAACACACCATTAAAGAAAAAGATTGGCGTGTGGATTAATGTCATAAGATT 3068
OY 354 ----- 354
Db 3069 GCGCGTATCGTATTACTGCACACTTTTGGATTAATTTAAAAAGCTTGGCTTAAGTAT 3128
OY 355 -----ThcGlyThrGlnTyrArgPheAspGluValPhePheThrIleAspPro 371
    ||||| :||| :|||
Db 3129 GCGACTAAGCGTGTATTCTCTATCTATGAGGATATT-----ATCAGCGCA 3176
OY 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArg-----GluLeu 389
    ||||| :||| :|||
Db 3177 AAAGACAAGCAAAAAATGGTGGAAAAAGCCAAAGTAGAGTTAAAAAATCCAGCAACAA 3236
OY 390 LeuGlnGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAla 409
    ||||| :||| :|||
Db 3237 TACGATCAAGGCGTCTCCTCAGCCAA---GAGCGTTAC----- 3272
OY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPhe 429
    ||||| :||| :|||
Db 3273 -----AATAGATCATTTGACACT-----TGACCTGAAGCAATGACAAATAGATAA 3320
OY 430 ProGluArgGlnGlnIleGlnAsnAspGlnValSerPheGlnGln----- 444
    ||||| :||| :|||
Db 3321 GAAATGATGACCGCTATCGCGCAAGATAAAGAGGCTTTACTCTATTTATATGATGCA 3380
OY 445 SerSerSerSerArgThrGlnProAlaGlnVal----- 455
    ||||| :||| :|||
Db 3381 GATAGCGCGGAGGGGATGCGCGCAAAATCCGTCAGCTTTCAGCGATGAGGGTCTT 3440
OY 456 -----AspGlnSerThrLeuGln---ProValIleGlnThrValGlnLeuThr 470
    ||||| :||| :|||
Db 3441 ATGACAACCGCGAGCGGAGATGATTCGAACGCCCATTTTCTAACTTTAAAGAGGGG 3500
OY 471 AspGlyIleLeuMetAspIleSerProIleGlnPheSerAlaSerAlaLeuIleGlnAsp 490
    ||||| :||| :|||
Db 3501 TTGAATCTCTTGAATCTCAATTCACAGCATGCGCTGAAGAGGCTTAGCGGATACA 3560
OY 491 LysLeuAsnLeuValAlaAla-----LysAlaArgHisLeuTyrAspMetProAspAsp 508
    ||||| :||| :|||
Db 3561 GCGCTAAAAACAGCCAAATGCGGGGATATTGACAGAAAGCTCATTTGTT---TCCCAA 3617
OY 509 ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528
    ||||| :||| :|||
Db 3618 AATGTCAAAGCTGTGCTGATGATTCGCGCACCATCAAGGAT---GAATTCAGC 3671
OY 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu----- 544
    ||||| :||| :|||
Db 3672 GAT---ATTCCGGTGGGAGTGAAGCTTGAACCTTGAAGAGCGCTATTTTGGGCGC 3728
OY 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
    ||||| :||| :|||
Db 3729 GTTTTATTAGACATGTGATCCCATTTAGCAATGAATCTTGTATTCGCGACACT 3788
OY 562 ProAlaAspValTyrGlnSerLysLysVal-----ProLeu 573
    ||||| :||| :|||
Db 3789 TTGATTGATCAAGAGGCTGCTAAAGGCTTGGAAGCCGGGATTAATTCATTAGCATC 3848
OY 574 TyrValPheValAlaSerAspLysProAlaGsp-----GlyGlnIleGly 588
    ||||| :||| :|||
Db 3849 CGCACCCCACTAACTTGAAGCGCAAGGCGTGTGCGCAAAATGCTATGCTTGAAT 3908
OY 589 LeuGlyTTPGly----- 592
    ||||| :||| :|||
Db 3909 TTGGGGAAGGCAAGATGACTTATCCGGGTGAAGCGGTGGCGTGTAGCCCGCAATCT 3968

```

```

OY 593 -----SerAspThrGlyThrArgLeuValThrLysPheGlnHis----- 605
    ||||| :||| :|||
Db 3969 ATTTGGGAGCCCTGGAACGACGCTCACTTAAAGACTTTCATGCTGGGGGACAGCGAGC 4028
OY 606 -----AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGly 618
    ||||| :||| :|||
Db 4029 AGGACTCAGGATGAGCGGGAATCTGACGAGCAAGAAAGGTTTGTGCGTTTACAAAC 4088
OY 619 LeuArgLeuSerGluAspLysGlyValLysLeuTyrAlaThrLys----- 634
    ||||| :||| :|||
Db 4089 CTTAGGACTTACACGAATTAAGAGGTAAGAAACATTATCGCTAACCGCGCTTACT 4148
OY 635 -----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeu 648
    ||||| :||| :|||
Db 4149 ATTTAGTGTAGAGCCCTAAGATTAAAGCGCTTTGATGGGGAATTGCAATTGAACG 4208
OY 649 GlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThr 668
    ||||| :||| :|||
Db 4209 GTTTATGAAGAAGTGTGTGAGC----- 4232
OY 669 LeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThrTyrSer 688
    ||||| :||| :|||
Db 4233 -----GTGAATAATGGCGATCAAGAGCTAATTTGTT 4265
OY 689 LeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrThrProGlnAspLeu 708
    ||||| :||| :|||
Db 4266 TTAAGG---ACAACGCAATATGTCAAG-----CCAACGCA 4298
OY 709 ProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu-----AlaGly 725
    ||||| :||| :|||
Db 4299 ---TTACCGCGCGTGGCGGATGAGGGAAGTGTATTTGCCCTTATGCTAGTGGG 4355
OY 726 ValAlaValHisLys-----ThrValAlaAspAsnLeuValAsnPrometArgGlyTyr 743
    ||||| :||| :|||
Db 4356 CATAGAGTGATTAAGGGGGAAGATATCGTATATTATC-----CAAGAGGCTGG 4406
OY 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
    ||||| :||| :|||
Db 4407 AATGTGCTATATGCAATCCCTTATGCGAGCAATTCCTAGTCAAGAT---AATGACCT 4463
OY 764 IleAlaArg-----AlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 780
    ||||| :||| :|||
Db 4464 ATTCCGCAAGATGTGTATGCCAAAGAAAGCGCTAATTCAAATATC-----TAT 4511
OY 781 GlySerAsnArgAlaHis---GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSer 799
    ||||| :||| :|||
Db 4512 GTTTAGAGGCTTAACCATTTAGAGCGGACCCATGGGATCAAAAAGCGCATATGTCAGT 4571
OY 800 AspAsnPheAsnHisValProTyrArgLeuArgPhePheAla-----Gly 814
    ||||| :||| :|||
Db 4572 GAAAA---GCGTTGTTTCCGGTATAGCTGATGATTAAT 4607
OY 815 GlyAspGlnSerIleArgGlyTyr---AlaHisAspSerLeuSerProIleSerAspLys 833
    ||||| :||| :|||
Db 4608 GGTAGGAAGACCCCTGCGCATTAATATGCGTAGGGGTTGTGAGTCTTGATGATTAAT 4667
OY 834 GlyTyrLeuThrGlyGlnValLeuAlaValAlaGlyThrAlaGlu----- 848
    ||||| :||| :|||
Db 4668 AGTGAAGTAGACATAATAGCGGATTTCAAAACGACGCTAACCTTTCAAAAGCATT 4727
OY 849 -----TyrAsnTyrGlnPheMetLysAspLeuArgLeuAlaValPheGly 863
    ||||| :||| :|||
Db 4728 GCCACATGGGATCCTTAACAACCCCTATATTCGCGACTTTAAA-----GCT 4775
OY 864 AspIleGly 866
    ||||| :||| :|||
Db 4776 AAGTGTGCT 4784

```

RESULT 30  
 US-09-815-242-7313  
 ; Sequence 7313, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:



APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
PRIORITY FILING DATE: 2001-03-21  
PRIORITY APPLICATION NUMBER: 60/191,078  
PRIORITY FILING DATE: 2000-03-21  
PRIORITY APPLICATION NUMBER: 60/206,848  
PRIORITY FILING DATE: 2000-05-23  
PRIORITY APPLICATION NUMBER: 60/207,727  
PRIORITY FILING DATE: 2000-05-26  
PRIORITY APPLICATION NUMBER: 60/242,578  
PRIORITY FILING DATE: 2000-10-23  
PRIORITY APPLICATION NUMBER: 60/253,625  
PRIORITY FILING DATE: 2000-11-27  
PRIORITY APPLICATION NUMBER: 60/257,931  
PRIORITY FILING DATE: 2000-12-22  
PRIORITY APPLICATION NUMBER: 60/269,308  
PRIORITY FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 7313  
LENGTH: 8673  
TYPE: DNA  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(8673)  
US-09-815-242-7313

## Alignment Scores:

Pred. No.:	0.0398	Length:	8673
Score:	127.50	Matches:	193
Percent Similarity:	33.90%	Conservative:	147
Best Local Similarity:	19.24%	Mismatches:	354
Query Match:	2.70%	Indels:	309
		Gaps:	52

US-09-914-168-2 (1-919) x US-09-815-242-7313 (1-8673)

QY 55 GlysnpProPovalleuLeuThrProgluInleGlnAlaArgLeuAsnAlaIagly 74  
DB 5023 GAGCGCCGCAAAATCATGTGCGCAATGAAAAA----- 5055  
QY 75 LeuAsnAlaLysProGlnSerGlnAlaLeuAspValAlaAspAspGlnSerPro 94  
DB 5056 -----ACGATGCTGCAGAAACCGTGATGTCTTTTGAATACGGCGAGCACT 5106  
QY 95 IleSerArgIleGlyGlnSerProProleuGlyLeuAspMetSerValIleGluIn 114  
DB 5107 AATCGCGTTAAAGGCGCTTAACAAAGCCCTTA----- 5139  
QY 115 ThrThrProLeuSerLeuGluLeuPheAlaGlnGlnSerThrGluMetGlyIleAsn 134  
DB 5140 -----AAATCGCTCAGTCAATCATTAAGGCAAGCGGCGCTTTCAGCAAAAC 5190  
QY 135 ProAsnAspTyrIleProGlnTyrGlnGlnGlnInProAsnSerGlnValAlaValPro 154  
DB 5191 CTTTAAAGTAAGCGCGGATTTTTCAGGC-----AGAAAGCTGATTTGCTTGGG 5241  
QY 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174  
DB 5242 CCTAATCTCAAAATGATGATGCGGCTTGCTTAATAACATGGCTTACACTCTTCAAA 5301  
QY 175 AspGlyValAlaAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly 194

DB 5302 -----CCGATTTGTATCCAAAGCT-----GAAGAGAGGC 5334  
QY 195 GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsn---IleLys 213  
DB 5335 TATGCCACCGCTCAAAACGAGCTTAACGATGATTAAGCAAAAGATGATGATG 5394  
QY 214 AlaAlaLeuGlnAspIleThrGlnGlnGlnSerAlaMetAspLeuAsnGlySerIleProArg 233  
DB 5395 GAGTCTTGCAAAATACAGAGGGGATCCGGTGCTACTACACCGCGCT---CCTACC 5451  
QY 234 LeuArgInThrAla-----LeuValAlaAlaArgAlaValGly 246  
DB 5452 TTGCACAAGCATTCATTCACGCTTCCATCCAAAGCTGATTTAGCGCAACATGATCA 5511  
QY 247 TyrTyr-----AspIleAsp----- 251  
DB 5512 TTGCACCGCTTAGTGTTACAGCTTCAAGCGGCAATTTTGAAGGGAGCCAAATGGCGGTG 5571  
QY 252 ---LeuSerIleIleArgAsnSerIleGlyGlu-----Val 262  
DB 5572 CATGTGCCCTTTAAGCCAGAACGATCGCTGAATGCAAGGTGCTGATGCTATG 5631  
QY 263 AspValIleIleHisAspLeuGlyLysProValTyrIleAsp----- 276  
DB 5632 AATATCCTTTTCCCTGCTAGCGTAAGCGGCTAGGCATTCCTAGCAAGATAGCTTTTA 5691  
QY 277 -----TyrArgAlaValGluValArgGlyGlyGluAlaAspAspLysAlaPheThr 293  
DB 5692 GGGCTTATATATCTTCTTTTAAGAAAGAGCGGCTCAAGCGGCGATAGCTTTTCT 5751  
QY 294 ThrValAlaAspGluValProLeuLeu---IleGlyAspValPheHisIleGlyLysTyr 312  
DB 5752 ACCGTGAATGAATCATCCGCCCATTCACAGAAAGATTTAGACATCCACCAAGAT 5811  
QY 313 GluThr-----LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAsp 330  
DB 5812 AGGTTTATAGATCAAGGCAATATATCCCTACAGATCA----- 5850  
QY 331 GlyArgTyrPheAspArgSerValAspValIleLeuProAsp----- 344  
DB 5851 GGGCGCATGATCATTTAAGTCC-----ATTTCGCTGATTTTATCCCTACGATTTGG 5901  
QY 345 -----AsnThrAlaAspValSerIleIleTyrAsp----- 354  
DB 5902 TCGAACAGACCCATGACAGAAAGATATTTGGCGTCTCTGTGATTAATGCTAAAGTT 5961  
QY 354 ----- 354  
DB 5962 GCGGTATCGATTAATCTGCAACCTTTTGTGAATTTAAAAACGCTTGCTTAGGTAT 6021  
QY 355 -----ThrGlyThrGlnTyrArgPheAspGluValAlaPheThrIleAspPro 371  
DB 6022 GCGACTAAGCGTGTGATTTTCTATCTCTATGAGAGATAT-----ATCACGCCA 6069  
QY 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArg-----GluLeu 389  
DB 6070 AAAGACAAACCAAAATAGTGAAGAAAGCAAGTATGAGCTTAATAAAATCCAGCAACAA 6129  
QY 390 LeuGlnGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAla 409  
DB 6130 TAGCATCAAGGCTGCTCACTACCAAA---GAGCGCTTAC----- 6165  
QY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429  
DB 6166 -----AATAGATCATTTGACT-----TGGACTGAAGTCAATGACAAATAGCTAAA 6213  
QY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln----- 444  
DB 6214 GAATGATAGCGCGATATCGCGCAAGATTAAGAGCGCTTAACTCTATTATATGATGCGCA 6273  
QY 445 SerSerSerArgTyrThrGluProAlaGlnVal----- 455

```

Db 6274 GATAGCGCGCAAGGGGTAGCGCGCGCAAAATCCGTCACCTTTTCAGCGATGAGGGTCTT 6333
OY 456 -----AspGluSerThrLeuGlu-----ProValIleGluThrValGluLeuThr 470
Db 6334 ATGACAAACCGCGGAGCGGATCATTTGAACGCCCATTTATTTCAACTTTAAACAGGGG 6393
OY 471 AspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAsp 490
Db 6394 TTGAATGCTTAAAGTAACCTTCATTCACCGCATGCGCTAGAAAGCGCTTAGCGGATACA 6453
OY 491 LysLeuAsnLeuValAlaAla-----LysAlaArgHisLeuTyrAspMetProAspAsp 508
Db 6454 GCGGCTAAACACCGCAATGCGGGGATTTTGACACGAAAGCTCATGATGTTT-----TCGCA 6510
OY 509 ArgValIleAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGluArgIleSer 528
Db 6511 AATGCAAGCGTGCTGTCTGATGATTCGCGCACCATGAAAGGATTT-----GAATACAG 6564
OY 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu----- 544
Db 6565 GAT-----ATTGGGCTGGGAGGAGCTGATTCGACCTTTAGAAAGCGATTTTGGCGCG 6621
OY 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
Db 6622 GTTTATTAGAAAGATGTGATGATCCCATTCGAATGAATCTTGCTTTATGCGGACACT 6681
OY 562 ProAlaAspValTyrGlnSerLysVal-----ProLeu 573
Db 6682 TTGATGATGAAAGGGGTGCTAAAGGTGTGAAACCGGAGATTAATCATTCATACATC 6741
OY 574 TyrValPheValAlaSerAspLysProArgAsp-----GlyGlnIleGly 588
Db 6742 CGCACCCAGTAACCTGTAAGCCCAAGGGCGTGTCCGGAATCTTATGCTTGTAAT 6801
OY 589 LeuGlyTrrGly----- 592
Db 6802 TTGGCGCAAGGCAAGATGATATCCGCGTCAAGCGGTGGCGTGTAGCGCGCAATCT 6861
OY 593 -----SerAspThrGlyThrArgLeuValThrLysPheGluHis----- 605
Db 6862 ATTGGGAGCGCTGGAACGACGCTCACTTTAAGGACTTTCCATGTGGCGGAGCAGCGAGC 6921
OY 606 -----AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlu 618
Db 6922 AGCAGTCACGATGAGCGCAATCTGACGACGAAAGAGTTTGTGCGCTTTTACAAAC 6881
OY 619 LeuArgLeuSerGlnAspLysGlyValLysLeuTyrAlaThrLys----- 634
Db 6982 CTTAAGACTTACACGAATGAAGGGTAAACACATTATGCTAACCGCGTAAACGCTTCT 7041
OY 635 -----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeu 648
Db 7042 ATTGTAGCTGTAAGCCTTAAGCATTAAGGCCCTTTTGATGGGAATTTAGCATTCGAACG 7101
OY 649 GlyTyrGlnGlnGluValPheGluHisSerThrAsnGlyPheAspLeuSerThrArgThr 668
Db 7102 GTTTATGAAGAGTCTGCTGAC----- 7125
OY 669 LeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyTyrPheAsnArgThrTyrSer 688
Db 7126 -----GTCAAAATAGGCGATCAAGAAAGCTAAATTTGT 7158
OY 689 LeuAlaGlyTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpIleAspLeu 708
Db 7159 TTAAAG-----ACAAGCGATATCTCAAG-----CCAAGCGAA----- 7191
OY 709 ProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu-----AlaGly 725
Db 7192 -----TTAGCCCGCGCTTGGCGGTAAAGATTGAGGGGAAAGTGTATTGCTTATGCTAGG 7248
OY 726 ValAlaValHisLys-----ThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
Db 7249 CATAGCGTGCATTAAGGGGGAAGTATCCCTGATTTATTATC-----CAAGAGCGCTGG 7299

```

```

OY 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 7300 AATGTGCTAATCCGATCCCTTATGCGAGCAATGTGATGCAAGAT-----AATGACCTT 7356
OY 764 IleAlaArg-----AlaGlyIleSerGlyValIleThrPheGlnGlyAspAsnAlaTyr 780
Db 7357 ATTCGCAAGCTGTGTATGCGCAAGAAAGAAAGCGCTTAATCAAAATAC-----TAT 7404
OY 781 GlySerAsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSer 799
Db 7405 GTTTAGAGGCTAACCATTTAGACGCGACCATGCGATGCAAAAAGGCGATATGGTGAGT 7464
OY 800 AspAsnPheAsnHisValProTyrArgLeuArgPheAla-----Gly 814
Db 7465 GAAAA-----GCGTTGTTCGCGTATGATGATGAT 7500
OY 815 GlyAspGlnSerIleArgGlyTyr-----AlaHisAspSerLeuSerProIleSerAspLys 833
Db 7501 GTTAGGGAACCGCTCCCATTTATGCTAGGGGTTCTGAGATCTTGATTTGATGATTAAT 7560
OY 834 GlyTyrLeuThrGlyGlnValLeuAlaValGlyThrAlaGlu----- 848
Db 7561 AGTGAAGTGAGCACTAATAGCGTATTTAAACCCAGCTAACACTTCAAAAGCAT 7620
OY 849 -----TyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
Db 7621 GCCACATGCGATCTTACACACCCTTATCATTTGCGGACTTTAA-----GCT 7668
OY 864 AspIleGly 866
Db 7669 AAGGTGGT 7677

RESULT 31
US-09-914-300-1524
; Sequence 1524, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berk, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085, 500-US
; CURRENT APPLICATION NUMBER: US/09/974, 300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680, 598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279, 526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1524
; LENGTH: 2545
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-914-300-1524

Alignment Scores:
Pred. No.: 0.00744 Length: 2545
Score: 127.00 Matches: 168
Percent Similarity: 35.72% Conservative: 161
Best Local Similarity: 18.24% Mismatches: 338
Query Match: 2.69% Indels: 255
DB: Gaps: 44

US-09-914-168-2 (1-919) x US-09-974-300-1524 (1-2545)
OY 65 GlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu 84
Db 260 CAGAGCAATTTGCGCTTATGCGCGAGC-----TGCTCGACGATATATCAAAAGTATG 316
OY 85 AspValAlaAsnPheAspAspGlnSerProIleSerArgIleGlyGlnGlnSerProPro 104

```

Db 317 AGCTTGAA-AACACGAGCGCGATGATATATCGGACCCTTGCCAAATACGCGAAAA 375  
Qy LeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGluLeu 124  
Db 376 GACGGATTGTAAGTGAAGATTGTTCCGCGACAGACACCTGACA-----CAGCTG 426  
Qy 125 AlaGlnIleuSerThrGluMetGlyIleAsnProAsnAsp-----Tyr 138  
Db 427 GCTACAGACGAGACGAGCGTCCGATCACAAAAGGCGATTACGATGCGAATACAT 486  
Qy 139 IleProGluTyrGlnGlyGluGlnProAsnSerGluValValProThrLeuGlu 158  
Db 487 ACAGCTGACACATGTCGCGGAAAAGTACGGA-----CTTACC 522  
Qy 159 ProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsn 178  
Db 523 CCTAACACAGATCATGATGATGAAA-----GGCTGATGGGAGACTCGTCACAG 570  
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyLysThrAla 198  
Db 571 AATATCCAGCGGTGCGCGGA-----GTCGCGGAAAAGACGCG 609  
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218  
Db 610 ATC-----AACCTTTGAAACAGATCCATACAGAGAGCTGCTGCTCA 657  
Qy 219 IleThrGlnIleuSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238  
Db 658 ATACACCAAGTCACGCGGAAAAAACTGAAGAAAGCTTGAAGATTAAAAAGACGCT 717  
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleLeuArgAsnSer 258  
Db 718 TTGATGAGCAAA-----GACCTTGCGACGATACGACAGAA 753  
Qy 259 IleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg 278  
Db 754 ---GCGCGCTTGAAATCAGCTCGATTCCTCGGATAT----- 789  
Qy 279 AlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGlu 298  
Db 790 -----GAAGGTTTGACAGCGGAGGAGCGGTCGTCAACATTTTAAACAT 831  
Qy 299 ValProLeuLeuIleGlyAspValPheHisIleGlyLysTyrGluThrLysLysAsnLeu 318  
Db 832 CTCGCG-----TTTCAT-----TCTTTG 849  
Qy 319 IleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAsnArgSerVal 338  
Db 850 CTCGAAAAGGATTGGCGAAGAGCGGAGAGAAAGAAAGAACAGTGGAA---GAGATC 906  
Qy 339 AspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGln 358  
Db 907 GATTCATGATC-----AAAAACAGACATC----- 930  
Qy 359 TyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThr 378  
Db 931 -----ACGATGACATGTTTTCG 948  
Qy 379 AspProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMet 398  
Db 949 TCACCGGCTTCCTCGTGTGTAAGACAG-----CTC 978  
Qy 399 GlyGluLysArgLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
Db 979 GCGACACATTTCACGAGCGCCGATCCTCGGT----- 1011  
Qy 419 TyrPheAsnMetValAsnThrGluIleValPhe-----ProGluArgGluGlnIleGln 436  
Db 1012 ---TTTTCGATCGTCAACGAGACGCGCTTCTTCATTCGGAAGAAGACGCGCTTCAG 1068  
Qy 437 AsnAspGlnValSerPheGlnGlnSerSerSerSerArgThrGluProAlaGlnValAsp 456  
Db 1069 TCGGAC-----TGTTTTAAAGAAATGGCGCGAAGATGAGTGAAGAAAAGAAATGGGTGTTT 1122

Qy 457 GluSerThrLeuGluProVal-----IleGluThrValGluLeuThr 470  
Db 1123 GATGCAAAAACGCGCGCGCTGCTTGGCGTGGCGGCGCATTAAGCTGAAGCCCTGAG 1182  
Qy 471 AspGlyIleLeuMetAsp-----IleSerPro-----IleGluPheSerAla 484  
Db 1183 TTTGACGTTCTTCTTGCGGCGCTATATCATCAACCGGGGCACTCCAGCATATGCGGA 1242  
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504  
Db 1243 AGCGTTGCGAAGAACATCAATATACATCGTTTCAGCGAGAT---GAAGCGCTTACGGA 1299  
Qy 505 -----MetProAspAspArgValLeuAlaIleAsnHisAspGly 518  
Db 1300 AAGAGCGGAAACACAGCGGCTTCGATGAAAGAGACTTGCC---GACCAT----- 1347  
Qy 519 ValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIle 538  
Db 1348 -----TTAGCCGAAAGCGAAGCGGATATCGCTGCTGCGCGGAAAAGCTC 1392  
Qy 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558  
Db 1393 CTGATGAACTGGAAGAAACAGCAGCTCGAAGT----- 1428  
Qy 559 ArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAla 578  
Db 1429 -----TTGACAGCGCTCGAAGATGCCCTGCGCCACATTTCTCGGT 1467  
Qy 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPglySerAspThrGlyThrArg 598  
Db 1468 GAGATGGAATCGATCGGTGCTCCAAAGTCGACAGACTGAAGAAAATGCGTGAAGAA 1527  
Qy 599 LeuValThrLys-----PheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla 615  
Db 1528 CTGTCCGCTAAATTAGCGGAATACGAGAAAAAATCCACAGC-----TCAGCA 1575  
Qy 616 GlyAlaGluLeuArgLeuSerGluAspLysLys---GlyValLysLeuTyrAlaThrLys 634  
Db 1576 GGTGAACCTTCACATTAATATTCACCAAAAGCAGCTCGGTGATTTGTT-----GAT 1629  
Qy 635 ProLeuSerHisProLeuAsnAspGlnLeuAlaThrLeuGlyTyrGlnGlnGluVal 654  
Db 1630 AAACCTCGGGCTGCTGCTCAAAAACCAAAAGC----- 1665  
Qy 655 PheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSer 674  
Db 1666 ---GGCTATTCACATTTCAGCCGATGCTTGAAAAGCTGGCGGATTAACATGTCATTATC 1722  
Qy 675 ArgSerIleIle-----GlnAsnGlyGlyTyrPasnArgThrTyr----- 687  
Db 1723 GAAGATATCTCTGACTACAGGAGATGTGGCAAGCTCCAGCTCACAATACCTAGAGGCTT 1782  
Qy 688 -----SerLeuArgTyrArgLeuAspLysLeuLysThr 698  
Db 1783 TTGAAGTGATCAAAAAGACAGCCATTAAGCGCATATACCCTTTAATCAACCTTAACA 1842  
Qy 699 Gln-----AlaProProGluThrThrPrgGlnAspLeuProValAspPheVal 713  
Db 1843 CAGACGGGAAGGCTCAGCTCAACAGCATCGGAATCTGCAAAATATACCGATCGCCTTGA 1902  
Qy 714 AsnGlyLys-----ProSerGlnGluAla----- 721  
Db 1903 GAGGCGCCGCAAAATCCGCGAGCGTTTGCTTGTGCGAAAAGGCGCTGCTCATTTTGG 1962  
Qy 722 -----LeuLeuAlaGlyValAlaValAlaLysLysThrVal 732  
Db 1963 GCGGATTAATCTACAGATCGAGCTGAGGTTTCGCGCACATTTCAAAGGATTAATAATTG 2022  
Qy 733 AlaAspAsnLeuValAsnProMetLargGlyTyrArgGlnArgTyrSerLeuGluVal--- 751  
Db 2023 ATTGAAGGTTCTCACGAAAGATATG---GACGTTTCAACACAAAACGCGATGAGCCTTTT 2079



```

Db 2241 TCTGCTGATAAAATATATAGTAGAGGAATTTCCACAGTATTTATCTCTCTACAAACA 2300
      :|||:||||| :|||:|||||
Qy 399 GYLUALATYRASNLEUGLALVALARGALALEUSERASNPLEULEALATHRAY 418
      :|||:||||| :|||:|||||
Db 2301 GAAATGACACATGAGAGAAAGAAACAGAACACATTAATCCAGAGTGAACAGACTACT 2360
      :|||:||||| :|||:|||||
Qy 419 TYRPHASNMEVLASNTHRGULLEVAL---PHEPROGLUARGULUGLNLLEGLNASN 437
      :|||:||||| :|||:|||||
Db 2361 TATACAGATGAAATTCACAGAAAGAGATCAATAAGTCCATTATGGAAGAAACAGAGAA 2420
      :|||:||||| :|||:|||||
Qy 438 ASPLGLNALSERPHEGLUGLINSERSESRERATGTHGLUPROLAGLINALSPGLU 457
      :|||:||||| :|||:|||||
Db 2421 GAAGCTTCTCTGGAGTGAACCTCTCATCTCTCTCAGACCAATTCATCTTACAGAG 2480
      :|||:||||| :|||:|||||
Qy 458 SERTHLEUGLUPROVALILEGLURHVALIGLULEUTHASPGLYLEULEUETASPLE 477
      :|||:||||| :|||:|||||
Db 2481 TCT-----TCTGTGAAATGACCAC----- 2501
      :|||:||||| :|||:|||||
Qy 478 SERPROILEGLUPHESERASLASERASNLEULLEGLNASPLYSLEUASNLEUVALA 497
      :|||:||||| :|||:|||||
Db 2502 -----TCTTTGATTTCCACACATGTGATA-----ACAAAGTTAAGTGA 2540
      :|||:||||| :|||:|||||
Qy 498 LYSALATRGHISLEUTHASPMETPROASPARGLVALLU-----ALALLEAN 514
      :|||:||||| :|||:|||||
Db 2541 GAGCCACAGAAAGTAAAGATGTGAGGAAAGACTTTACAGAACTCCAGCTACTACMAAA 2600
      :|||:||||| :|||:|||||
Qy 515 HISASPARSPGLVALASNARGSERILEU-----GLYARGILESERASPLALVALSER 532
      :|||:||||| :|||:|||||
Db 2601 TATGTGAAATATATTACACACAGCTTTTGCCCATGTGACTTTTANGTCTTGACAGACC 2660
      :|||:||||| :|||:|||||
Qy 533 ALAVALLAARGALALEUETHLEUPROASPSGLUSERGLUASNGULVALILEASPLEUPROGLU 552
      :|||:||||| :|||:|||||
Db 2661 ACTGTATCAAAATGCTCATG---GATGAAGATTAATACAACTCCAAAGCTTTAGAGTCT 2717
      :|||:||||| :|||:|||||
Qy 553 ARGTHRALALEUALASNARGLYSTHRPROALASPVALLTYR----- 566
      :|||:||||| :|||:|||||
Db 2718 ACAGAACCTTGACCTCTTCAAAATGTCCTGCTTACTCACAACCTGTGGGATGAAT 2777
      :|||:||||| :|||:|||||
Qy 567 ---GINSERLYLSVALPROLEUTHRYVALPHEVALALASERASPLYS----- 581
      :|||:||||| :|||:|||||
Db 2778 GGAAGAGATTAAGACATCCCAAGTTTCACTGAAGATGGAGCAGATGAATTTACTCTTAT 2837
      :|||:||||| :|||:|||||
Qy 582 PROATGASPSGLYGLNLLEGLYLEUGLYTRPGLYSERASPTHGLYTHRGULEVALTHR 601
      :|||:||||| :|||:|||||
Db 2838 CCAGATAGTACTCAAAACAGTGAAGAGGTTACTGATGAGACATACACCCCATGGA 2897
      :|||:||||| :|||:|||||
Qy 602 LYSRHE-----GLUHISASNLEULEASNARGASPSGLYTYRGLNALAGLYLAGULEU 619
      :|||:||||| :|||:|||||
Db 2898 AAATTCACAATTAGATTTCAGCCAACTACATCACTGGTATTGCAAGAAAGTCAACTTGG 2957
      :|||:||||| :|||:|||||
Qy 620 ARGLEUSERGLUASPLYSGLYVAL-----LYSLEUTHYR 631
      :|||:||||| :|||:|||||
Db 2958 AGAGATTCTACACACTGCAAGAAAGTTCACCTATCACAGACACTGAAAGCCAACTTAT 3017
      :|||:||||| :|||:|||||
Qy 632 ALATHRYSPROLEUSERHISPROLEUASNAPSGLNUARGALATHRGULEGLYTYRGLN 651
      :|||:||||| :|||:|||||
Db 3018 CCAACCATGGAAGAAAT-----GCTTTGGGTGAAGTA 3050
      :|||:||||| :|||:|||||
Qy 652 GLINGULVAL-----PHEGLYHISSETHRASN 660
      :|||:||||| :|||:|||||
Db 3051 GAAGATGTGACCTCTCTAAGCCATCTACTGTTCCTCCCAATTTGCACACACTTCAGAG 3110
      :|||:||||| :|||:|||||
Qy 661 -----GLYPHEASPLEUSERTHRYATGTHRGULU----- 670
      :|||:||||| :|||:|||||
Db 3111 GTGGAAGATTAAGCATTTTGTATTAGTATAGCACCCCAAGACCTTACTATTATGATGAC 3170
      :|||:||||| :|||:|||||
Qy 671 -----HISGLULISERATSERILEILEGLNASNGLYLYTRPASNARGTHRYTSER 688
      :|||:||||| :|||:|||||
Db 3171 TCTTCCATACCATCTCTCTTCTGTAATTCACAAAGACAGACTGGGA----- 3218
      :|||:||||| :|||:|||||
Qy 689 LEUATGTYRARGLEUASNPSLYSTHRYGULNALAIPROGLUATHRYTRPGLINASPLEU 708
      :|||:||||| :|||:|||||

```

```

Db 3219 -----CTGTAGTACTCTCTGTCCATCAGAA----- 3245
Qy 709 PROVALASPPHEVALASNLYLSPROSERGLINGULALEULEUVALAGLYVALAVAL 728
      :|||:||||| :|||:|||||
Db 3246 -----GATGAAGTCTTAGTGAACCTCTCAAGACATCTGTCTC-----ATT 3287
      :|||:||||| :|||:|||||
Qy 729 HISLYSTHRVALALASPASNLEUVALASNPRO 739
      :|||:||||| :|||:|||||
Db 3288 GATCAGACTCGCCTTGACAGCACTATTCTCCA 3320
      :|||:||||| :|||:|||||

RESULT 33
US-09-954-531-1380
; Sequence 1380, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: weaver, zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 1380
; LENGTH: 9534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-1380

Alignment Scores:
Pred. No.: 0.0916 Length: 9534
Score: 124.50 Matches: 198
Percent Similarity: 34.64% Conservative: 147
Best Local Similarity: 19.88% Mismatches: 351
Query Match: 2.63% Indels: 301
DB: 9 Gaps: 50

US-09-914-168-2 (1-919) x US-09-954-531-1380 (1-9534)
Qy 12 PHEMETPROVALALALEUVALATYRLEUPROLEUMET----- 24
      :|||:||||| :|||:|||||
Db 4470 TTTCTTCGCGTAACAATTTCACGCC-----CCTCTGTGTGCGAGAAAGACTTGACGACT 4523
      :|||:||||| :|||:|||||
Qy 25 THRSERGLNALALEUVALGLINGLNASNAPROVALASNILLEASNHISVALPROVALA 44
      :|||:||||| :|||:|||||
Db 4524 ACCGCGCACGGCTGTCCACGGGGATATGAAGCCGCTGACTGTGAAGGTGGCCCTG 4583
      :|||:||||| :|||:|||||
Qy 45 HISASPTHRALALLEASN-----GLNALALYLAGLYASNPRO----- 57
      :|||:||||| :|||:|||||
Db 4584 GCTATAGTGGCAGTCCAGCAACCTCGAGGCTCTGCCAAGAAATGTGATGATCCCT 4643
      :|||:||||| :|||:|||||
Qy 58 -----PROVALLEUTHRPROGLUGLNLLEGLNALATRGLEUASN 71
      :|||:||||| :|||:|||||
Db 4644 ATGGCTCACTGGCTGTGGCCTGTGACCCGTGCACAGATTTGCCAGTCCGCACTGGAG 4703
      :|||:||||| :|||:|||||
Qy 72 -----ALALAGLYLEUASNALALYSPROGLINSERGLN 82
      :|||:||||| :|||:|||||
Db 4704 CCACGGGAGAAAGTGTGACGGCTGCACACCTGCATCGACCGCAGG----- 4751
      :|||:||||| :|||:|||||
Qy 83 ALALEUASPVALLASNPHASPARSPGLINSERPROLISERATGILEGLYGLUINSER 102
      :|||:||||| :|||:|||||
Db 4751 ----- 4751

```

Oy 103 ProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 122  
 Db 4752 -----GCTGGAGTGTCTTTTGTGTGGA-GATGAGTCACTGGCTTCTCGGTGAC 4804  
 Oy 123 LeuPheIaGInGluSerThrGluMetGlyLeuAsn-----ProAsn 137  
 Db 4805 TTGGCTGCCCTGGACAGACATGCTGATGACATCAACCTCACTGCTCCGCTCGCGCA 4864  
 Oy 138 TyrIleProGluTyrGInGluGluProAsnSerGluVal-----ValValProPro 155  
 Db 4865 TATAAATCGTGTATGCTTGAATAATGATGATCAGAGCACTAAAGCACTTCTGCACT 4924  
 Oy 156 ThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 175  
 Db 4925 CAGCGGCCCGCAGAGAG-----CTTATTCACTGGCAGAGGCAATCTG---AATACA 4975  
 Oy 176 GlyValAsnLysValProArgLeuLysAlaLysPheTyrGInSerSerGInSerGlu 195  
 Db 4976 CTCGTGACCGAAATGAAAGAGAGCTGTCAGCCAGGCTACCAAGTGCAGCAGATGGCGAG 5035  
 Oy 196 ThrSerAlaIleGlySerSerHisGInLysThrGluProTyrAlaAsnIleLysAla 215  
 Db 5036 CAGACG-----GGACAGGATGCTGAGAGAGCAACACAAAGAGCAAG-----TCC 5080  
 Oy 216 LeuGluAspIleThrGInGluSerAlaMetAsp-----LeuAsnGlySerIlePro 232  
 Db 5081 CTGGGAAATTCATTCAATAGAGAGCTGCCCGGATGCAGAGCTGTAAATGAAAGAACTATA 5140  
 Oy 233 ArgLeuArgGInThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAsp 252  
 Db 5141 AACTTAATGAAACT-----CTAGCAACTGCA-----GACGAG 5173  
 Oy 253 SerIleIleArgAsnSerIleGly-----GluValAspValIleIleHisAsp 269  
 Db 5174 GCCCTTGACGAAATTTTGAAAGGCTTCAGAAAGAGATGACCAAGATTAAGAACTG 5233  
 Oy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGluGluLysAlaAsp 289  
 Db 5234 AGGAGG-----AAAACTAGACAGACACAAAAAGAAATTCCTGGAAGAT 5275  
 Oy 290 Lys-----AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPhe 307  
 Db 5276 GACTGTGATCATCGAGAAAGCCCTTCGAAAAAAAGTGAAGAGCTGTTCGAGAGTCCGG 5335  
 Oy 308 HisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnHisAspValHisGly 327  
 Db 5336 GGGGAAATGAA---GAATGGAAGAGATCTCGGAAAAAAGCTGCTGACTCAAAAAC 5392  
 Oy 328 TyrPheAspGlyArgTyr-----LeuAspArgSerValAspValIleLeuProAspAsn 345  
 Db 5393 AAGTTGATGATCTTGGGAGCTTTTGAGAGAACCCACAGATTAATCAAGAGCAAT 5452  
 Oy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGInTyrArgPheAspGluVal 365  
 Db 5453 CGC----- 5455  
 Oy 366 PhePheThrIleAspProLysThrAsnGluLeuThrThrAspProAspLysLeuProVal 385  
 Db 5456 CTAATTGGCACTAATACAGAA-----AACATGACTGCATTGGAGAAA 5497  
 Oy 386 LysArgGluLeuLeuGluGluLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlu 405  
 Db 5498 AAGAGAGCGCTGTGAGAC-----CGCAACGCA---CAAAATTGAG 5536  
 Oy 406 AlaValAlaLeuSerAsnAspLeuIle-----AlaThrArgTyrPheAsnMetVal 423  
 Db 5537 AACACTTTAAAGAGAGCAATGACATACATCGATGAAGCCAGCTCTTCAGATGAATC 5596  
 Oy 424 AsnThrGluIleValPheProGluArgGluGluGIn----- 436  
 Db 5597 AACTCATATAGACTATGTGTAAGACATCCAAACTAAATTGCCACCTATGCTCGAGAG 5656  
 Oy 437 ---AsnAspGluValSer-----PheGluGlu 444

Db 5657 CTTATATGATAAATATGATGACCTCTCCCAAGAAATTAAGACAGAAAGCTTGACAGAG 5716  
 Oy 445 SerSerSerArgTyrGluProAlaGluValAspGluSerThrLeuGluProValIle 464  
 Db 5717 GTGTCCAGGCTGAGCCAGCAGCTCAGTGTGAATGACATCATCT----- 5761  
 Oy 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484  
 Db 5762 -----GCTGTCTTGATGGAATCTTGATGAGCTTAAGAAACATCTCTTCAATGCC 5812  
 Oy 485 -----SerAsnLeuIleGlu-----AspLysLeuAsnLeuVal 495  
 Db 5813 ACTGACGCTTCAAAAGCTTACAGCAATATTAAGACATATATGATGAAGCTGAGAAAGTT 5872  
 Oy 496 AlaAlaLysAlaArgHisLeuTyrAsp----- 504  
 Db 5873 GCCAAAGAGCCAAAGATCTTGACATGAAGCTTCAAAAGCTGCAACAGCTCTCGGGGT 5932  
 Oy 505 ---MetProAspAsp-----ArgValLeu----- 511  
 Db 5933 TTATTAAAGAGATGCGCAAAAGCTGTCTTCAGAAAGCTTCAGGATTTCAAGCAAGCC 5992  
 Oy 512 -----Ala 512  
 Db 5993 AAGAGTTACCAATGATGTAAAGAAATGAAGACCATCTAAATGCTTAAACCAGG 6052  
 Oy 513 IleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAlaValSer 532  
 Db 6053 ATGAAATATGCTGATGCTAGAAATGGGAGATCTCTTGAAACTTTGAATGACACTTTGGGA 6112  
 Oy 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552  
 Db 6113 AAGTTA-----TCAGTATTCCAAATGATACAGCTGTAAAGCTGCAAGCTGTAAAGGAC 6166  
 Oy 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGInSerLysValPro 572  
 Db 6167 AAGGCCAGACAGCCAAAC---GACACAGCTAAAGATGTACTG---GCACGATTAACAGAG 6220  
 Oy 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGInIleGlyLeuGlyTyrGly 592  
 Db 6221 CTCACACAGAACTCGATGCTGCGTGAAGAAAGATTAACAATAACTAGCA-----GAC 6271  
 Oy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIleAsnArgAsp 611  
 Db 6272 AGCGTGCCCAAAAGCAATGTGCTGTAAAGATCTTCCAAAGAAACAAATCATGTCCGAT 6331  
 Oy 612 GlyTyrGluAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyr 631  
 Db 6332 GCAGATGCCACTGTCAAAAATTTAGAACAGGAAGCTGACCGCTAATAGATTAACTC--- 6388  
 Oy 632 AlaThrLysProLeuSerHisProLeuAsnAspGluLeuArgAlaThrLeuGlyTyrGlu 651  
 Db 6389 -----AAACCCATCAAGGAA---CTTGAGGATACCTTAAGAAAAACATCTGTGAGATA 6439  
 Oy 652 GInGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671  
 Db 6440 AAGGAATTTGATAACCAACGCT-----CGAACAACAGCCAAAT 6475  
 Oy 672 GluIleSerArgSerIleIleGInAsnGlyGlyTyrAsnArgThrTyrSerLeuArgTyr 691  
 Db 6476 TGTATCAAGATATCTGTCTTCAGAGAGTGACTGCTATTCGACATAC----- 6523  
 Oy 692 ArgLeuAspLysLeuLysThrGlnAlaProProGluThrTyrGInAspLeuProValAsp 711  
 Db 6524 -----AAACAGCAAAATCAAGAAAGAGATTTACATATATTTGTGTGCAAC 6568  
 Oy 712 PheValAsnGlyLysProSerGInGluAlaLeuLeuAlaGlyValAlaHisLysThr 731  
 Db 6569 -----GTAAGACAGCT 6580  
 Oy 732 ValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGluArgTyrSerLeuGluVal 751

Dh 6581 GTTGCCTAATCACTCTC-----TTTATCTT 6607  
Qy 752 glyserSerGlyLeuValSerAspAlaHisMetAlaIleAlaArgAlaGlyIleSerGly 771  
Db 6608 GGAAGTGGCAAAATTATTACTTCTGCGCTATAGAAATGCGTAAGGCAAGTCCAGCTTC 6667  
Qy 772 ValTyrSerPheGly-----AspAsn 778  
Db 6668 CTCGGAGTGTGGATCTGGAGTTGGAGCTGTAGATCCAGATTGTGACTATTATATAC 6727  
Qy 779 AlaTyrGlySerAsnAlaGlyAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrp 798  
Db 6728 TCATTTGTGATCCGTATGTCATCAAGAACTGG-----AGAAATGAACTATT-- 6778  
Qy 799 SerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspSerIle 818  
Db 6779 -----TCGTGAGAGCCCTGGATGAGCCCAAGCCAGC 6811  
Qy 819 IleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspGlyTyr----- 835  
Db 6812 ATTCGCGCCAGCACACACATTCGACGCTCTCT-----CCAGGATACAGATTCTA 6862  
Qy 836 ---LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMet 854  
Db 6863 GATGTGATGCAATGCAATGCTGTGTTGTGTGGCTGACTGGGAATTAAAGAAAGCT 6922  
Qy 855 LysAspLeuArgLeuAlaValPhe---GlyAspIleGlyAsnAlaTyr 869  
Db 6923 GATCTGTACTGTGATTTACATCTACTGCTGATGGAGAAACATAC 6970  
RESULT 34  
US-09-782-378A-23  
; Sequence 23, Application US/09782378A  
; Patent No. US20020102731A1  
; GENERAL INFORMATION:  
; APPLICANT: Hearing, Patrick  
; APPLICANT: Bahou, Nadie  
; APPLICANT: Sandalon, Ziv  
; APPLICANT: Gatenko, Dmitri  
; TITLE OF INVENTION: Adenoviral Vectors  
; FILE REFERENCE: STONYB-04970  
; CURRENT APPLICATION NUMBER: US/09/782,378A  
; CURRENT FILING DATE: 2001-02-12  
; PRIOR APPLICATION NUMBER: 60/237,747  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 23  
; LENGTH: 10302  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-782-378A-23  
Alignment Scores:  
Pred. No.: 0.103 Length: 10302  
Score: 124.50 Matches: 138  
Percent Similarity: 37.35% Conservative: 119  
Best Local Similarity: 20.06% Mismatches: 258  
Query Match: 2.63% Indels: 173  
DB: 10 Gaps: 36  
US-09-914-168-2 (1-919) x US-09-782-378A-23 (1-10302)  
Qy 146 GlnProAsnSerGluValValProProThrLeuGlnProGluLysProGlyLeuIle 165  
Db 5440 CAAAGAGAGAGAAATGTTTACATCAACCTTATGAGAGATTAATAAAAAAGATC--- 5496  
Qy 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla 185  
Db 5497 ---CGTTTGCAATATTATTCTTTCATCTAGATACAAATAT-----AAGCA 5544  
Qy 186 LysPheTyrGlnSerSerGlnSerGlyLys---ThrSerAlaIleGlySerSerHisGln 204  
Db 186 LysPheTyrGlnSerSerGlnSerGlyLys---ThrSerAlaIleGlySerSerHisGln 204

Dh 5545 ATCCCTATTCAACAGAGAAATGGTCAACTTGCTTGTGATTTAGATCATCACTTCT 5604  
Qy 205 LysThrGluProTyrAlaAsnIleLysAlaAlaLeu-----GlnSplIleThrGln 221  
Db 5605 CCTACAGATTATCTGGTTGAATTTAAACAAATTTTACTTTGATGATGATGTT----- 5658  
Qy 222 GluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAla 241  
Db 5659 -----GANTTATCGCTTAATGTTCAGAGCTC---AACACTGCTATTTC--- 5700  
Qy 242 AlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg-----AsnSer 258  
Db 5701 ---GAAGACTTCTCTTTTCAGAGAACCTCTGAGAAATATACAAAGCCACTGGACAA 5757  
Qy 255 IleGlyGlnValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyr 277  
Db 5758 CTYGGAGAGCAGATTGTCATCTATCTATGAAAAACAGCCAGATGTCATCTTGAAGCTCT 5817  
Qy 278 -----ArgAlaValAlaGluValArgGly----- 284  
Db 5818 GGACCTGAAGCCATTTCAGATCAGATACATCTACTACAGCTGAATGCAAAATGGACAGA 5877  
Qy 285 -----GluGlyAlaAspAspLysAla----- 291  
Db 5878 ATTAATAGATGTACAGTATCGAAGAGGTGTTTTCAGAGGCAATGGAGAAATGGAGA 5937  
Qy 292 -----PheThrThrValAlaAspGluValPro 300  
Db 5938 CAGTTCCATTGTGACCTTAATGACCTCACACAGTGGATNACAGAGCTGAGAA----- 5991  
Qy 301 LeuLeuIleGlyAspValPheHisHisGlyLysTyrGlyThrLysLysAsnLeuIleGlu 320  
Db 5992 TTACTGGTTGATACCTGCTGCTCCAGTGGCAGCTCGATTTAGAGAAACCCAGATACAT 6051  
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArg-----TrpLeuAspArg 336  
Db 6052 CAGCAGAACTTGAGGTGGCTGATCAGACCCAGCCAGCTTTTTCAGACCTTAACCGCA 6111  
Qy 337 SerValAspValIleLeuProAsp---AsnThrAlaAspValSerIleTyrAsp--- 354  
Db 6112 ACTGGGAGTGGATTTGTGAGAACTCTCCAGCAGATGAGAGCTTCTTGAAGCAAAA 6171  
Qy 355 ---ThrGlyThrGlnTyrArgPheAspGluValPhePheThrIleAspProLysThr 373  
Db 6172 CTGGCAGGTTTAACCAACGCTGCGATGCATTTGTT---GCAGAGTGAAGATAGGACAG 6228  
Qy 374 AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlnLeu 393  
Db 6229 CCAAGGCTAAGAGAGAAAGTAAGCAGGTGATGAGATCAGCAGCTTACATGAGATT 6288  
Qy 394 Leu-----ThrValAspMetGly 399  
Db 6289 ATCTGTTGTTAACAAAGGCTGACATGCTATGCAAAAGATCAACACCAATTTGGCA 6348  
Qy 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla----- 416  
Db 6349 GAA-----AACCTCAGAAATTAAGAGCTTAACCAAGAAATGGAATACATGCTGAA 6402  
Qy 417 ---ThrArgTyrPheAsnMetValAsnThrGluIleVal-----Phe 429  
Db 6403 AAACCTCAATGCTGATGAATGAACCTGAATGATGATCTTCAGATAAAGTCTGATTGA 6462  
Qy 430 ProGluArgGluGlnIleGlnAsnAsp-----GlnValSerPheGluGlnSerSer 447  
Db 6463 CCGTGAAGGATAAATTTTCAGAAAGCTTAAGGACTGTAAATATGACATGATTAAGATT 6522  
Qy 448 SerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrVal 467  
Db 6523 TGCACAGAGGTGCTTACACACCTTGAAGAAATCATCCAGAGACC----- 6567  
Qy 468 GluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu 487  
Db 6568 -----AGTTCTGTTTCACAGACAGAGATTTGCTCATCTCAAT 6606

```

OY 488 IIEGLIASPLYSLEUASNLEUVALAIALALYSALARGHISLEUTYRASPMECPRO--- 506
DB 6607 GTCCAA---AAGCTGTGCTAGTATCATCTCGCA-----GATATTCCTGT 6651
OY 507 ---ASAPAPARGVALLEUALAIIASNHISASAPSGLYVALASNARGSERILEUGLY 525
DB 6652 CAGTCTCATCTGACTTCGGAAATTTCAATTCCTGCTGATCTTGATTAACATTAACA--- 6708
OY 526 ARGILESERASP-----AlaValSerAlaValAlaAlaValAlaValProAspGlu 542
DB 6709 GAACCTACCCGACTGCTGATTAATGACACAGATGCTGAAGTCCAACTTGCTACACTGT 6768
OY 543 SERGIUNSGIUNVALIIEASPLEUPROGLUARGTHRALAVALASNARGLYSTHPR 562
DB 6769 GCGGATGTAGACAGATCAAT---AAGACCTTCCCGAATGAAATTAACAAAG 6819
OY 563 ALASAPVALTYRGINSERLYSVALPROLEUTYRPHLEVALALASERAPLYSPRO 582
DB 6820 GCTGACTTAGACACGCCCTCCAGCTGATTAATTTTACATTTGGCAGCAAAATTTG 6879
OY 583 ARGASPCILGIIIEGLYLEUGLYTYRPGLYSERASPTHREGLYTHARGLEUVALTHRLYS 602
DB 6880 AAAAAT-----AAAGCTCCAGTTCCAGATATGGAACACCATTTACAGAAAA 6927
OY 603 PHEGLUHIISANLEULIASNARGSPGLYTYRGINALGLYALAGLULEUARGLEUSER 622
DB 6928 TTGGAACGCTCAAGAACCCAGTGGATGGCACCACCATGGCGTTGAGCTAAGACAGCAG 6987
OY 623 ---GLUASP----- 624
DB 6988 CAGCTTAGACACGTGATTTAGACAGCTTTCAGTGGATGACCATAGGAGAGACTGAA 7047
OY 625 -----LysLysGLYValLYSLEUTYRALATHRLYS-----Pro 635
DB 7048 GAACGTATGAGAAATATGAGGCTGACCTATATCTTCCACCAAGCCGACGGATCCA 7107
OY 636 LEUSERHISPROLEUASNAPOLN-----LeuARGALATHRLLEUGLYTYRGINGLNU 653
DB 7108 CTCACCAACAAATTTCTGATTAACCAATACTGCTTCAAGAACTGGGCTCGAGATGCT 7167
OY 654 VALPHEGLYHISERTHASNGLYPHEASPLEUSERTHARGTHRLLEUGLHIISGLUTLE 673
DB 7168 ATC-----GTCATGGCGTTCAT-----AACGTC 7191
OY 674 SERARGSERILEIEGLIASNGLYLYTRPASNARGTHRYRSEULEARGTYRARGLEU 693
DB 7192 CTGCAAGAACCTCGAGAGATATGAGGAGTGTACACAAAGATGTGAAGAACACACA 7251
OY 694 ASPLYSLEUTYRTHGINALAPROPROGLUHTHTRPGLINSPLLEUPROVALASPEHEVAL 713
DB 7252 GAGTACTTAAAAACA-----TCATGATCAATCTCAAAACAAAGTATTGCT 7296
OY 714 ASNGLYSPROSERGLINGUALA 721
DB 7297 GACAGACAGAACGCTTGAGGCT 7320

RESULT 35
US-10-108-605-102
; Sequence 102, Application US/10108605
; Patent NO. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 311338
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16

```

```

; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 102
; LENGTH: 14155
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; US-10-108-605-102

Alignment Scores:
Pred. No.: 0.163 Length: 14155
Score: 124.50 Matches: 181
Percent Similarity: 33.04% Conservative: 116
Best Local Similarity: 20.00% Mismatches: 347
Query Match: 2.63% Gaps: 259
DB: 9 Indels: 41

US-09-914-168-2 (1-919) x US-10-108-605-102 (1-14155)

OY 110 SERVALIIEGLIUNTHRTHRPROLEUSERLEUGLULEUPHEALINGLUSERTHR 129
DB 8934 AGTGCTCGAGAGACTTAACAGGTACACCTGGCTTTCACACGTCAAAATTTGAA 8993
OY 130 GLUMETGLIASNPROASNAPTYRILEPROGLUTYRGINGLINGLINSER 149
DB 8994 TATTACGACCCAGCTGGCCGATGAACCTGAGCCCAAGGTTAACTATTGATCCCAACAG 9053
OY 150 GLUVALVALPROPTHRLLEUGLUPROGLULYSPROGLYUULEYLSARGLEUTYR 169
DB 9054 GTGGACCTGAGTCCATCC-----AAAAAGCCCAAC 9083
OY 170 ALARGLEUPHEASNSPGLY-----ValASNLYSVALPROARGLEUYSALALYS 186
DB 9084 AGCGAATTTGAATTCGATGTCGCAAGTCTTATGCCCAACAGGTCAACACAGCCTGGCAAC 9143
OY 187 PHEYRGLINSERSEGLINSERGLIUNTHR 197
DB 9144 GCCCTGATATCCGTGAGGATGAGCACCACTTTGGTAACATTAACAGTTCATTATGAT 9203
OY 198 ---AlaIIEGLYSERSERHISGLINLSTYRGLUPROTYRALASNLIELYS----- 213
DB 9204 GAGCGCGTCAAGAGTGCAGACAGCCAAAGAGCCATGCTTCGGTGAAGCTTTCC 9263
OY 214 -----AlaAlaLEUGLUSPILETHGLN 221
DB 9264 AAGAAATCTGAGCGGAGCGACGACAAAGATGATGCTGCTGAGCAAA---GCACAG 9320
OY 222 GLUSERALAMEASPLEUASNGLYSERILEPROARGLEUARGLINTHRLAILEUVALA 241
DB 9321 CACATTTTGGGCACAATACGAGCAACGATTCGAACTTACACCCATTAACAAAGTCTG 9380
OY 242 ALARGALAVALLGLYTYRASP----- 249
DB 9381 GAAAGGCGAGGAAACCTATGACAAAGTCAACCTAGTTCGCCATCAAGCGCA 9440
OY 250 ---IIEASPLEUSERILEIARGASNSERIIEGLYGLUVALASPVALLIIEIHISASP 268
DB 9441 AACAGAGACCTCAATGCCCTCAAGAAATGATTTGAGAAATTCAGCGACCATCTGGAAGAT 9500
OY 269 LEU-----GLYLUPROVALTYRILEASPTYRARGALAVALLGLUVALARGLY--- 284
DB 9501 CTTTCAACTGAGCGAAGCTTCACAAAGCCAGCTCGCTGACCGCCCAATGTG 9560
OY 285 -----GLIUGLYLAASAPLYSALAPHEPTHRYVALAIALASPGLUVALPROLEU 301
DB 9561 GCCAACAGAGAGCTTTGACAACTTAATTCGACACCGTTTCGAGCAA 9611
OY 302 LEULIEGLYSPVALPHENISHISGLYLYSTYRGILTHRLYSLEUINLEUGLUSN 321
DB 9612 -----AAGCTACAGCGAGAGAAC----- 9632
OY 322 ALASERVALIIEGLIUNTHRTHRPHASPGLYARGTYRLEUASNARGSERVALASPVALLIE 341

```



Db 9632 ----- 9632  
 Qy 342 LeuProaspasnThrAlaaspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPhe 361  
 Db 9633 -----ATCAAGATGCTGGCAAT----- 9650  
 Qy 362 AspGluValAlaPhePheThrIle-----AspProLysThrAsnGlnLeuThrThrAsp 379  
 Db 9651 -----TTCCCTCATCAATGGCGATCTGACCCCTGCAATCAATTAATACAGAG 9695  
 Qy 380 ProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGly 399  
 Db 9696 TTGGATAACTTG-----AGAGATGCTTAACGAGCTCAATTCCTCATATAAA--- 9743  
 Qy 400 GluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr 419  
 Db 9743 ----- 9743  
 Qy 420 PheAsnMetValAsnThrGluIleValPheProGluArgGlnGln----- 434  
 Db 9744 -----AATGTCGACGAGAATTTACCCCTGAGGGAGACACCATAGAGGCG 9791  
 Qy 435 -----IleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThr 450  
 Db 9792 GATGCCCTTACCGATCAAGCCGACAGAGGCGGCGGATTTAGCCATTAAGCCCAAGAT 9851  
 Qy 451 GluProAlaGlnValAspGlnSerThr-----LeuGluProValIleGluThrValGlu 468  
 Db 9852 CTGGCTCCCAATACAGGAGATGATGTCGAGTGGCGGCGGATCAAGCCCAACT 9911  
 Qy 469 LeuThrAspGlyIleLeu-----MetAspIleSerProIleGluPhe 482  
 Db 9912 GCCACTACGGGAATCGTTGAGCGCGTTGAACTGCGCGAAGACATGACAGATGCATC 9971  
 Qy 483 SerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeu 502  
 Db 9972 TCTGCGCGCGAATGACACTGATAGACATGATGATGAAGAGAGCT---CATCTC 10028  
 Qy 503 TyrAsp-----MetPro 506  
 Db 10029 GCTGACACTGGATCTACTGATCTTCTCCAGACAGCCAGCTCTGCGAAGGTACA 10088  
 Qy 507 AspAspArgValLeuAlaIleAsnHisAspArgly-----Val 519  
 Db 10089 GAGACACTAGAGCCCCCTGTAACGCCCTGGCGGCAAGGTCCAGAAATCTCACCCCTG 10148  
 Qy 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539  
 Db 10149 AACAAATCCACCGAACACACAGCTAAGATATTAAACAGCTGATGACCAA---CTG 10202  
 Qy 540 ProAspGlnSerGluAsnGlnValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559  
 Db 10203 CCAGCTAGTGTGCGAGAGGATATG-----TGGAAGAATTTCAATGCCAAT--- 10247  
 Qy 560 LysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPhe-----Val 577  
 Db 10248 -----GCCAGTGAATGCTGTGAGATACTGAAGAACCTGCTAGAGATCTGGAACCTGTG 10301  
 Qy 578 AlaSerAspLysProAlaGspGlyGlnIleGlyLeuGlyTyrPglYSerAspThrGlyThr 597  
 Db 10302 AGTGTCCAAAGCCAAAGAGACTGCGAAGGACACATGGCATAGAGATTGG----- 10355  
 Qy 598 ArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAla 617  
 Db 10356 -----GATTTGACCAATTAAGGCGTTTCTCAAGCTACAA 10391  
 Qy 618 GluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSer 637  
 Db 10392 CAATTG-----GATGACGTAGAAGCA---TCCGTTTGAAGCTAAAGCAATTTGCA 10439  
 Qy 638 HisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlyHis 657

Db 10440 GAGATATCGAGAACAG---CAGCACCGTGGGCGACTGACAGCCGCGCATTTGGCCAG 10496  
 Qy 658 SerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIle 677  
 Db 10497 GAATTCGAGAACTTGAAGGCCCAAGTGAAGCT---GCTCGGCAATTTGGCCAAAGCATC 10553  
 Qy 678 IleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeu---AspLysLeu 696  
 Db 10554 AAGGGGGCGCTCAATTTCAAGCCAAAGCAGATCTGGAAGTGAAGACACCGGAAAGAC 10613  
 Qy 697 LysThrGlnAlaProProGluThrThrPglAsnAspLeuProValAspPheValAsnGlyLys 716  
 Db 10614 AAGTTGCTAGCC-----ACTGCAACATATATGACCTATTTCGGACCCAGC 10664  
 Qy 717 ProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLysThrValAspAsn--- 735  
 Db 10665 CCATCCGTTTCCCTTTTATCTGCGC---AATGCAACACAGCCGCGGAGAAAC 10721  
 Qy 736 -----LeuValAsnProMetArgGlyTyrArg---GlnArgTyrSerLeuGlnVal 751  
 Db 10722 GACTTTTGCGCCGTTGAATTTGTAATGCTATGCGATTTCCATAGATTGGCAAT 10781  
 Qy 752 GlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGly 771  
 Db 10782 GGGCGGAACCATCACCAGCGATGATGACGATGCGCATGCGCTGATTCACGCTGT 10841  
 Qy 772 ValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly 791  
 Db 10842 GTGATCGCATGATGCGCAATGCGCAAGCTAACATATTAGGAAGAACTAACCCAGCGTAC 10901  
 Qy 792 Ile-----GlnAlaGlyTyrIle---TyrSerAspAsnPheAsnHisValPro 806  
 Db 10902 GTGCTTGACAGACAGCAAGCTGCGCTATCTTGAAGATCCCAAGATATCTTCATGTGAT 10961  
 Qy 807 TyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyr---AlaHisAsp 825  
 Db 10962 AAGAACACTGCCCTATTTGTTGCTGATCCGCGGTATTTGCGATCAATGCTCCGCCA 11021  
 Qy 826 SerLeuSerProIleSerAspLysGlyTyrLeuThrGlnGlnValLeuValAlaGly 845  
 Db 11022 GACTTGACCACTACTCTTCTCCGCGACATGAGATCTTAAGATTTGGCGATGACAC 11081  
 Qy 846 ThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIle 865  
 Db 11082 GTAGATTTGTGAACCTT-----GTATATGCGCAT--- 11111  
 Qy 866 GlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAla-----Gly 882  
 Db 11112 -----GATAACGACCAAGGCGCGCGAGCGCGGAT 11141  
 Qy 883 ValGlyValArgTyrPheLysSerProValGlyGlnValArgValAspValAlaThrGlyVal 902  
 Db 11142 GTGCTGCTGAGAGACAGAACGCGGTA-----ACTGCTGTG 11177  
 Qy 903 LysGlnGlnGlyAsn 907  
 Db 11178 CGGTTCAGAGGTAAAT 11192  
 RESULT 36  
 ; US-09-808-880-1  
 ; Sequence 1, Application US/09808880  
 ; Publication No. US20030027287A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Belach, Mary C.  
 ; APPLICANT: Shah, Sanjay Krishnakant  
 ; APPLICANT: McDaniel, Robert  
 ; APPLICANT: Tang, Li  
 ; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE  
 ; FILE REFERENCE: 30062-20029.00  
 ; CURRENT APPLICATION NUMBER: US/09/808,880  
 ; CURRENT FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: US/09/428,517  
 ; PRIOR FILING DATE: 1999-10-28



Db 22602 AGTGGCGTGGCGGACATCTCTTTTCGACGACGTCGACGAGCCGACGCGCCGCGGCGCTG 22661  
 Oy 574 TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySer 593  
 Db 22662 TCCACGGCTGGCGACCAACGAC-----GGACAGGCCCGGAGCAG 22700  
 Oy 594 AspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspLysTrp 613  
 Db 22701 GACGGCGGTACGACACTGCTAGACAGCTCG----- 22730  
 Oy 614 GlnAlaGlyAlaGluLeuArgLeuSerGlnAspLysLysGlyValLysLeuTyrAlaThr 633  
 Db 22730 ----- 22730  
 Oy 634 LysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlu 653  
 Db 22731 CPTCCGCTGACCGAG-----ACGAGACAAACGACGAGACTGCTGTCTGCTCCAGAGT 22784  
 Oy 654 -----ValPheGlyHisSerThrAsn----- 660  
 Db 22785 GAATGCTGGCGCTGACCTGACCTGCTCCACCGACGCGGTCCACGACCGCGCTTC 22844  
 Oy 661 -----GlyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIle 677  
 Db 22845 CGAGAGATCGGGTGTGACTGACTGACACGCGTC----- 22877  
 Oy 678 IleGlnAsnGlyLysTrpAsnArgThrLysSerLeuArgTyrArgLeuAspLysLeuLys 697  
 Db 22878 -----CAGCTCCGG 22886  
 Oy 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro 717  
 Db 22887 AACCGGCTTACGGCCGACACGAGGCGCATGGCGCTTCGACAGACGCTGTTCGACTACCG 22946  
 Oy 718 SerGlnGluAlaLeu-----LeuAlaGlyValAla-ValHisLys 730  
 Db 22947 ACCACCAACGAGACTGCGCGAGTACTGCGGTCCGAACACTGTTCGCTGTCCGCGCACCA 23006  
 Oy 730 SThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGlu 750  
 Db 23007 G-----CTGACCTCTCCGCTGCTCCGGAACGCGGAGAGGAGACGACGCC 23051  
 Oy 750 uValGlySerSerGlyValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSe 770  
 Db 23052 GTCCGTCATCGTGGGATGGCTGCGGTCCCGGCGGAGGTGATACGCGGAGACCTTC 23111  
 Oy 770 rGlyValTyrSer-PheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrG 790  
 Db 23112 TGGAGCTGTCTGGAAGCGG-----CGGCGATGTCATCTCCAACTT 23153  
 Oy 790 LysGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808  
 Db 23154 CCGGCCA-ACCGCGGCTGGGACATGGACACTCTGAACCCGACCCGAGCGGA 23208  
 RESULT 37  
 US-10-107-649-1  
 : Sequence 1, Application US/10107649  
 : Publication No. US20030044949A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Behr, Regine  
 : APPLICANT: Sloma, Alan  
 : TITLE OF INVENTION: Polypeptides Having Gamma-Glutamyl  
 : FILE REFERENCE: 10157-200-US  
 : CURRENT APPLICATION NUMBER: US/10/107,649  
 : PRIOR FILING DATE: 2002-03-27  
 : PRIOR APPLICATION NUMBER: 60/279,374  
 : NUMBER OF SEQ ID NOS: 6  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 1  
 : LENGTH: 1815  
 : TYPE: DNA

	US-10-107-649-1	ORGANISM: <i>Bacillus agardhaerens</i>
Alignment Scores:	0.00908	Length: 1815
Prod. No.:	124.00	Matches: 114
Score:	37.12%	Conservative: 92
Percent Similarity:	20.54%	Mismatch: 219
Best Local Similarity:	2.62%	Indels: 130
Query Match:	9	Gaps: 30
DB:		
US-09-914-168-2 (1-919) x US-10-107-649-1 (1-1815)		
QY 190	SerserGlnSerGlyGluThrSerAlaIleGlySerHisGlnLysThrGluProTyr	205
DB 163	AGCGAAATATATAGTCAACCGCGTAACAAATATATGCGAGCGCTGTCGATATTAT	222
QY 210	AlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMetAspLeu-----	227
DB 223	GGCAGACGACCTCCGC-----CATCCGTTACTCCAGAACTGGAAATGACATTATCAG	276
QY 228	AsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr	247
DB 277	AATGGCGGT-----ACTGCTATCGATCCCGCTGGCGTTCATT	318
QY 248	TyrAspIleAspLeuSerIleIleArg-----AsnSerIleGlyGluValAspVal	264
DB 319	ATG-----CTTAACGTAAGTTGACCTTACGGAATCTGGAAATGGAGCGCGGTG	363
QY 265	IleIle---HisAspLeuGlyGluProValTyr---IleAspTyrArg---AlaValGlu	281
DB 370	ATGCTCTACACGACCGACCGACGAAAGGGGTCTAATGATTATGATTATCGTAAGAGCGCCCT	429
QY 282	ValArgGlyGluGlyAlaAspAsp-----	288
DB 430	ATAAGCGGAAAT-----GATGACCTTACTGCGGAGGGGTGCATCCCTGCTTGT	483
QY 290	LysAlaPheThrThrValAlaAsp-----	297
DB 484	AAGGAATGACCTCATTCATGACAAATCATGAGAACTCCGTGGGAAGCTTATGCC	543
QY 298	-----GluValProLeuLeuIleGlyAspValPheHisGlyLys	311
DB 544	CTGCCTATGAAAGGGCGTAACAGGTTTTCAGTAGTGATATTTCACACCAACA	603
QY 312	TyrGluThrLysLysAsnLeu---IleGluAsnAlaSerAlaGluHisGlyTyrPheAsp	333
DB 604	GGTAAATGCTGTTGCTTACTTGAATGGAAACAGACGCGCTCACTATTCTTCCGAA	663
QY 331	GlyArgTrpLeuAspArgSerValAspAlaIleLeuProAspAsnThrAlaAspValSer	350
DB 664	GGACAGCGCTCTGGGTGTCATAGACCAACTCGTTCAAGAAGATTAGCAGATACGTTACGA	722
QY 351	LeuIleTyrAspThrGlyThr-----	357
DB 724	CTCATTTGAGGAAACCGGTGCGATGTTTACAGTGACCAATCGGGGATCTTTCAAA	783
QY 358	---GlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeu	376
DB 784	CACAAATTCACCTTACTGGAAGAAGATCTGGTACGTGACCAACAATTAAGTGA	840
QY 377	ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlnGluGlnIleLeuThrVal	366
DB 841	-----CCGTCTCTCTGACAGTGGAGAAACAATTTGCTATGCT	879
QY 397	AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla	416
DB 880	GGTCCCTACCTTCATCTGGAACAGTAGTGCTCAACACTTTG-----	921
QY 417	ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGlnGlnIleGln	436
DB 922	---CAATGGCAGATTCAGCTCGATTAAATGACCTTTTCCCGAAT---GAAAGACTTGGCC	975

```

Oy 437 AsnaspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGln----- 454
    |||
    |||
    |||
Db 976 GAAGAT-----TTCTCATTCGAGACCTTGTCTAACTCGAGATTACAGACAT 1026
Oy 455 -----ValAspGlnSerThrLeuGluProValIleGluThrVal 467
    |||
    |||
    |||
Db 1027 TATATTCATTGATMAATGAATAAATAACAAAGCCACTATGACGCCGCTTGACACT--- 1083
Oy 468 GluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu 487
    |||
    |||
    |||
Db 1084 ---TTAGCTGACCCCTGCGCTTGCATGATATTCATCATCAAGACATCAGATGATCATTTAT 1140
Oy 488 IleGlnAspLysLeuAsn-----LeuValAlaIleLysAlaArgHis 501
    |||
    |||
    |||
Db 1141 ATTCACCAATTACTAGACGATATTTCTTTATATGAGATTACCTGGTATACCTCCGAG 1200
Oy 502 LeuThrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg 521
    |||
    |||
    |||
Db 1201 TTATTCGATTTCTCTCGCAGAGAACCCGATTTCTGCACATACCCTCATTTGCTTATTTGTA 1260
Oy 522 SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArg-----Ala 537
    |||
    |||
    |||
Db 1261 GATAAAGAGGACGACATGCTATCCGCTACCATTCACATGATGTAATTTGGTTCTGCT 1320
Oy 538 IleLeuProAspGlu-----SerGlnAsnGluValIleAspLeuProGluArgThrAla 555
    |||
    |||
    |||
Db 1321 ATTTATATTTGACGGCTTTTATATCAATATCAATGACTAATTTTATGATGATATCCACAT 1380
Oy 556 LeuAlaAsnArgLysThrProAlaAspValIleThrLysSerLysValProLeuThrVal 575
    |||
    |||
    |||
Db 1381 TCATTTATGCTTTATGAGCCGAGTAAAGGCCGATCTTTTTCGACCCATG---ATT 1437
Oy 576 PheValAlaSerAspLysProArgAspGlyIleGlyLeuGlyThrProGlySerAspThr 595
    |||
    |||
    |||
Db 1438 TTT-----GAAGAAGAGGCCGACGATTCGCGGATGCGCTCACCAAGT 1482
Oy 596 GlyThrArgLeu-----ValThrLysPheGlnHisAsnLeuIle 608
    |||
    |||
    |||
Db 1483 GGAAGACGATATCTCTGCTATGCTATTTTCACAGCATTCATGCATATAT---CATATGAGATA 1599
Oy 609 AsnArgAspGly-----TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAsp 624
    |||
    |||
    |||
Db 1540 AATGACGATGCTGATCCATGACACTTCACAGAACGATCGAGGCTCCCGCTTTTATATAC 1599
Oy 625 LysLysGlyValIleLysLeuThrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeu 644
    |||
    |||
    |||
Db 1600 GAAGAAAGATGCT---ACTCTATTACAAAGAGGAATTACCTGAAGATGTTAGCAAGAACTT 1656
Oy 645 ArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThr 659
    |||
    |||
    |||
Db 1657 CCC---AATATGGGATAC-----TCTGTGTTGGACATAGCTCA 1692

RESULT 38
US-09-738-877-2
; Sequence 2, Application US/09738877
; Patent No. US20020015970A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan
; APPLICANT: Weiss, Stephen J.
; APPLICANT: Glynn, Richard
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND ME
; FILE REFERENCE: A-69806/DJB/JUD
; CURRENT APPLICATION NUMBER: US/09/738,877
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/US 00/22061
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 4
```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 5346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-877-2

Alignment Scores:
Pred. No.: 0.0441 Length: 5346
Score: 124.00 Matches: 180
Percent Similarity: 34.07% Conservative: 130
Best Local Similarity: 19.78% Mismatches: 301
Query Match: 2.62% Indels: 301
Gaps: 45

US-09-914-168-2 (1-919) x US-09-738-877-2 (1-5346)

Oy 31 GlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHis----- 45
    |||
    |||
    |||
Db 2825 CAGAAAGAAAGAACCCCGCCGCTTACTGAACT-CTGCCAGAAAGACGAGAGCCCGGGGC 2883
Oy 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGln 65
    |||
    |||
    |||
Db 2884 GACAGGCTGTTAGTGAAGCGGAA-----TTGACCCCGAAAGCT 2922
Oy 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
    |||
    |||
    |||
Db 2923 GTGACAGCTGCAGAAACCTCGAGGGCCATTTGGTTCCGAAAGAGACCGAGCATCTGCT 2982
Oy 86 ValValAsnPheAspAspGln---SerProIleSerArgIleGlyGluGlnSerProPro 104
    |||
    |||
    |||
Db 2983 GCTGAAGAGACCACAGCAAAATGTTGTACAGACTCTCCAGTTAACCGAC---TCCCA--- 3036
Oy 105 LeuGlyLeuAspMetSerValIleGluGlnThrThrProLeu----- 118
    |||
    |||
    |||
Db 3037 -----GACACCACAGAGAGGCCACTCCGCTGACGAGGTGAAGGTGCC 3081
Oy 119 -----SerLeuGluGln 122
    |||
    |||
    |||
Db 3082 GTACCTGACATAGAAGAGCAAGAGAGCGGACCTCAAGAGTCTCCAGGACGTGGCAGAA 3141
Oy 123 LeuPheAlaGlnGlnSerThrGluMetGlyLe---AsnProAsnAspThrIleProGlu 141
    |||
    |||
    |||
Db 3142 AAATGAAAGAGAAATCCAGCTGCTGGCACCGGTGGGCCAGAAAGATGTGCTTCAGCCT 3201
Oy 142 TyrGln-----GlyGluGlnProAsnSerGluValValIleProThrLeuGlu 158
    |||
    |||
    |||
Db 3202 GTCGAGAGAGCAGAGGCGAAGAACACAGAGACGAGCT----- 3240
Oy 159 ProGluLysProGlyLeuIleLysArgLeuThrAlaArgLeuPheAsnAspGlyValAsn 178
    |||
    |||
    |||
Db 3241 ---GAAGCGTGGGCTCGAAGAAAGAG-----ACGAGATGAGTGTG 3279
Oy 179 LysValProArgLeuLysAlaLys-----PheTyrGlnSerSerGlnSerGlyGlu 195
    |||
    |||
    |||
Db 3280 AAATAGATGCTCGAGAGGCAAAATGAGCTTTTACACAAAGGAAGTGTGGGCGAG 3399
Oy 196 ThrSer-----AlaIleGlySerSerHis 203
    |||
    |||
    |||
Db 3340 ACCACCCCGAAGACCTTTGAAGAAAGCTCTCAAGTCAACAGACAGACATAGAGTCCAGAG 3399
Oy 204 GlnLysThr-----GluProTyrAlaAsnIleLysAlaIleAlaLeuGluAspIle 219
    |||
    |||
    |||
Db 3400 CTTGTAACCACTTGTCAGAGCGGAAGCACTTAGCTGGGTAAATCA-----CAGGAGATG 3453
Oy 220 ThrGlnGlnSerAlaMet-----AspLeuAsn 228
    |||
    |||
    |||
Db 3454 GTGATGGAACAGGCTATCCCGCTGACTCGGTGGAAGACCCCTACAGACAGTGAAGTAT 3513
Oy 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaIleArgAlaValGlyTyrTyr 248
    |||
    |||
    |||
Db 3514 GGAAGCACACCC-----GTAGCCGACTTT 3557
```



PRIOR APPLICATION NUMBER: PCT/US 00/22061  
 PRIOR FILING DATE: 2000-08-11  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 1  
 LENGTH: 6608  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 PUBLICATION INFORMATION:  
 AUTHORS: Naert et al.  
 TITLE: Gravin, an autoantigen recognized by serum from myasthenia gravis  
 TITLE: Patients, is a kinase scaffold protein  
 JOURNAL: Curr. Biol.  
 VOLUME: 7  
 ISSUE: 1  
 DATE: 1997-01-01  
 PAGES: 52-62  
 DATABASE ACCESSION NUMBER: U81607  
 DATABASE ENTRY DATE: 1997-06-26  
 US-09-738-877-1

Alignment Scores:  
 Pred. No.: 0.0601 Length: 6608  
 Score: 124.00 Matches: 180  
 Percent Similarity: 34.07% Conservative: 130  
 Best Local Similarity: 19.78% Mismatches: 301  
 Query Match: 2.62% Indels: 301  
 DB: 10 Gaps: 45

US-09-914-168-2 (1-919) x US-09-738-877-1 (1-6608)  
 QY 31 GINGHNASNP... 45  
 DB 3016 CAGAAAGACACCCCGGCTACTGAACT-CTGCCAGACAGAGAGGCGCGGCGC 3074  
 QY 46 AAPTTHALIA... 65  
 DB 3075 GACAGCGTCGTAGTACGCGGAA- 3113  
 QY 66 IIEGLIALI... 85  
 DB 3114 GTGACACCTGCAGAAACTCGAGGCCATTGGCTCCGAGAGAGACCGCATCTGCT 3173  
 QY 86 VALVALA... 104  
 DB 3174 GCTGAGACACACAGAAATGCTGTACAGACTCTCCAGTTAACCGAC--TCCCA 3227  
 QY 105 LEUGLYLE... 118  
 DB 3228 -----GACACCACAGAGAGGCCACTCGGTGCAGAGGTGGAAGTGC 3272  
 QY 119 -----SerLeuGluGlu 122  
 DB 3273 GTACCTGACATAGAGAGCAGAGAGCGGACTCAAGAGTCTCTCCAGCAGTGCAGAA 3332  
 QY 123 LEUPHEAL... 141  
 DB 3333 AAGTGAAGAGAAATCTCCAGCTGCTGGACCGGTGGGCCAGAAATGTCTTACGCT 3392  
 QY 142 TYRGLN... 158  
 DB 3393 GTGCACAGACAGAGCCAGAAAGACCAAGAGAGAGGCT- 3431  
 QY 159 PROGLI... 178  
 DB 3432 ---GAAGCTCGGGTCTGAAAGAACAG- 3470  
 QY 179 LYSVALP... 195  
 DB 3471 AAGAGTAGTCTCAGAGGCAAAACTGAGCTTTACACAAAGGAGAGTGTGGGCGAG 3530  
 QY 196 THRSE... 203  
 DB -----AlaIIEGLYSerSerHis 203

DB 3531 ACCACCCAGAAAGCTTTGAAAAAGCTCTCAAGTCACAGAGAGCATAGTCCAGTGCAG 3590  
 QY 204 GINLYSTH...-GluProTyraIleAsnIleLysAlaIleLeuGluAspIle 219  
 DB 3591 CTGTAAACCATTTGTCACCGCAACCTTACCTGGGGTAAATCA-----CAGAGATG 3644  
 QY 220 THINGLU... 228  
 DB 3645 GTGATGGAACAGGCTATCCCTGCTGCTGAGAAACCTTACAGACAGTGAACATGAT 3704  
 QY 229 GLYSER... 248  
 DB 3705 GGAACACCC- 3728  
 QY 249 ASPLEA... 268  
 DB 3729 GACGACACAGCCACCAACCAAGAGAGATGTGTGAA-----ATCCATGAG 3776  
 QY 269 LEUGLY... 288  
 DB 3777 GAGATGAGTCT- 3818  
 QY 289 ASPLYA... 308  
 DB 3819 GCAGAGCAGTCTCTGACAGAAAGAGGCGCTCCAGACCTTCCAGTTTGTGTTCCAG 3878  
 QY 309 HISGLY... 328  
 DB 3879 GAA-----GAACTAAAGAACAAATCAAGATGGAAGACCTCTGTAGAGCAT 3923  
 QY 329 PHEASP... 348  
 DB 3924 -----ACAGTTAAAGAGGTGTGCTGCTG- 3950  
 QY 349 VALSER... 360  
 DB 3951 GTATCCATCTGTCAAAAGCTGAGGGACTCAAGAGCGTGCACCATATGCTGATGAGAA 4010  
 QY 361 PHEASGL... 375  
 DB 4011 ACCAAAGACGTACATTTTTCAGAGACTTGAGGGGTCTATAGAC-----ACAGC 4061  
 QY 376 LEUTH... 383  
 DB 4062 ATTAACAGTCACTGGGAAAGGTCACTGAGTGGCTTAAAGTGGAAGGACAGAAAGAA 4121  
 QY 383 ----- 383  
 DB 4122 GCTGAATGAAAAAGATGATGCTCTTGAACTGCAGAGTCAAGTCAAGTCTCCATCC 4181  
 QY 384 PROVAL... 394  
 DB 4182 CCGGTGAGAGAGAGATGTGTTCAAGTCAAAAGGAGAAACAGAGAGAGCCAAAC 4241  
 QY 395 THYVAL... 424  
 DB 4242 CATGTGAATGAAGAGAGCTTGACAGCAAGAGCTGTACCGTA-----TGTGAAGAG 4295  
 QY 415 ILEAL... 433  
 DB 4296 GTCACTAAGACGCTCTCCAGACAGTGAATGTCCCATCATATGATGGGCAAGAAAGTTC 4355  
 QY 433 ----- 433  
 DB 4356 AGCAGTTTGAAGAGAGCCCTCTCTGCTAGTCAAGAGAGGAGCATATGACCAAA 4415  
 QY 434 ---GLN... 452  
 DB 4416 ATTCAAGTTCAGAGCTCTGAGCATATCTTACACAGCGGCTGACAGAGAGGAGAAAG 4475  
 QY 453 ALAGN... 472  
 DB 4476 GTCTTAGAGAAACTGCCAAACATTTTAGAAACAGGTGAACGTTGAGCTGACAGTGCA 4535

```
QY 473 ILEuMeCAsPILeSerProILeGluPheSerAlaSerAsnLeuILeGlnAspLysLeu 492
    |||||
Db 4536 CATTAGTT-----CTGGAAGAGAAATCTCTGAAAAAATAAGAC----- 4577
QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTYrAspMetProAspAspArgValLeuAla 512
    |||||
Db 4578 -----TTGGCCGCTCAT-----CCAGGGGAGATGATCTGTCGCC 4610
QY 513 ILeAsnHisAspAspGlyValAsnArgSerILeLeuGlyArgILeSerAspAlaValSer 532
    |||||
Db 4611 ACAGGGCCCGAGCTGTGAGGCAAAATCGACACAGTG-----ATAGTATCT 4655
QY 533 AlaValAlaArgAlaILeLeuProAspGluSerGluAsnGluValILeAspLeuProGlu 552
    |||||
Db 4656 GCTACTACCAAGAAAGGCTTAAGTCCGACCTGGAAGAGAG-----AAA 4700
QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTYrGlnSer--LysLysVal 571
    |||||
Db 4701 ACCACATCAGTGAAGTGCAGATGAGATGACAGGCTTGCTGCAGAGAGTTC 4760
QY 572 ProLeuTYrValPheValAlaSerAspLysProArgAspGlyGlnILeGlyLeuGlyTTP 591
    |||||
Db 4761 AAAGTGAGTGTACAAATGAGGATTGAGCCTGAAATGGATTGGAACCTGAGACC 4820
QY 592 GlySer----- 593
Db 4821 AAAGCAGTAACCTGTCCAAACATCAGACAGCCGTTGACACAGTTGTAGCTACA 4880
QY 594 ---AspThrGlyTYrArgLeuValThr-----LysPheGlnHisAsnLeuLeasn 609
    |||||
Db 4881 GAAGAAACAGCCAGCAATGTTGACCTGTGACTTACAGACACAGCTGATGATAAAA 4940
QY 610 ArgAspGlyTYrGlnAlaGlyValGluLeuArgLeuSerGluAspLysGlyValLys 629
    |||||
Db 4941 GCTGACAGCCAGAGAGCTGAGAGAGAAAGCAAGAGAGAGAGAACCTGAGGCTCT 5000
QY 630 LeuTYrAlaThrLysProLeuSerHisProLeuAsnAspGluLeuArgAlaThrLeuGly 649
    |||||
Db 5001 GCACAGGATGAAACACCAATCTTCAAGCAAAAGAGAGACTCAGAGTCAAGCCAGTGGGA 5060
QY 650 TYrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeu-----SerThr 666
    |||||
Db 5061 CAA-----GCACATCTCATATTTCCAAAGACATGATGAAACCTCAGAA 5105
QY 667 ArgThrLeuGluHisGluLeuSerArgSerILeILeGlnAsn-GlyGlyTTPAsnArgH 686
    |||||
Db 5106 AAGACCATGACTGTGAGGTAGAGAGTTCCTCACTGTAATGATCAGCAGCTGGAAGAGGTC 5165
QY 686 TYrISer---LeuArgTYrArgLeuAspLysLeuLysThrGlnAlaProGluTYrThr 705
    |||||
Db 5166 GTCTCCCATCTGAGAGAGAGGAGTGGAGCTGGA-ACAAGCTCTGCTCCAGAA----- 5219
QY 705 pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaG 725
    |||||
Db 5220 -----GATGATGCTCATGCTCTTGTAGCA-- 5243
QY 725 yValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTYrArgG 745
    |||||
Db 5244 -----GAAAGATAGAGAGAGTCACTAGTGAACCGAAAGAGAT----- 5282
QY 745 nArgTYrSerLeuGluValGlySerSerGlyLeuValSerAsp-----AlaAsnne 762
    |||||
Db 5283 -----GAAAAAGGTGATGAT-----GTTGATGACCTGAAAAACCAAGACTC 5323
QY 762 tAlaAlaAlaArgAlaGlyLysSerGly 771
    |||||
Db 5324 AGCCCTGCTGATACTGATGCTCAGGA 5351
RESULT 40
US-09-880-107-3439
: Sequence 3439, Application US/09880107
: Patent No. US20020142981A1
```

```
GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 3439
: LENGTH: 6608
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 U81607
US-09-880-107-3439

Alignment Scores:
Pred. No.: 0.0601 Length: 6608
Score: 124.00 Matches: 180
Percent Similarity: 34.07% Conservative: 130
Best Local Similarity: 19.78% Mismatches: 301
Query Match: 2.62% Indels: 301
DB: 10 Gaps: 45

US-09-914-168-2 (1-919) x US-09-880-107-3439 (1-6608)
QY 31 GlnGlnAsnAspProAlaAsnILeLeasnHisValProAlaHis----- 45
    |||||
Db 3016 CAGAGAAAGAACCCCGACGGTTACTGAACT-CTGCCAGAGAACAGAGAGCCCGGGGC 3074
QY 46 AspThrAlaLeuAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrProGluGln 65
    |||||
Db 3075 GACACGGCTGTAGTACGAGCGGAA-----TTGACCCCGGAGACT 3113
QY 66 ILeGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
    |||||
Db 3114 GTGACAGCTGCAGAAACTCAGGCGCATGTGGTCCGAGAGAGAACCGAATCTGCT 3173
QY 86 ValValAsnPheAspAspLys-----SerProILeSerArgILeGlyLysLeuSerProPro 104
    |||||
Db 3174 GCTGAAGAGAGCAAGAAATGTTGTCAGAGTCTCCAGATTAAACGAC---TCGCCA--- 3227
QY 105 LeuGlyLeuAspMetSerValILeGluGluTYrThrProLeu-----SerLeuGluGlu 118
    |||||
Db 3228 -----CACACACAGAGAGAGGCCACCTCCGCTCAGAGAGTGAAGTGGC 3272
QY 119 -----SerLeuGluGlu 122
    |||||
Db 3273 GTACCTGACATAGAGAGAGAGAGCGGACTCAAGAGTCTCCAGCAGTGGCAGAA 3332
QY 123 LeuPheAlaGlnGluSerThrGluMetGlyIle---AsnProAsnAspTYrILeProGlu 141
    |||||
Db 3333 AAAGTGAAGAGGAATCCAGCTGCTGCGACCGGTGGCCAGAAATGTGCTTCAGCCT 3392
QY 142 TYrGln-----GlyGluGlnProAsnSerGluValValProThrThrLeuGlu 158
    |||||
Db 3393 GTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3431
QY 159 ProGluLysProGlyLeuLeuLysArgLeuTYrAlaArgLeuPheAsnAspGlyValAsn 178
    |||||
Db 3432 ---GAAGCGTGGGTCTGAGAGAGAG-----ACGATCTAGTGTG 3470
QY 179 LysValProArgLeuLysAlaLys-----PheTYrGlnSerSerGlnSerGlyGlu 195
    |||||
Db 3471 AAAGTAGATCTCAGAGAGCAAAACTGAGCCTTTTACACAGAGAGAGAGAGAGAGAGAG 3530
```

OY	196	ThrSer-----	-----AlaIleGlySerHis	203	
Db	3531	ACCACCCGAGAAGCCTTTGAAAAAGCTCTCAAGTCACAGAGAGCGCATTAAGCTCAGAGTG	3590		
OY	204	GlnIysThr-----	-----GluProTyrAlaAsnIleLysAlaIleLeuGluAspIle	219	
Db	3591	CTTGTAAACCACTTGTCAAGCGGAAAGCCTTAGCTGGGGTAAATATCA-----	CAGGAGATG	3644	
OY	220	ThrGlnIserAlaMet-----	-----AspLeuAsn	228	
Db	3645	GTGATGACGAAACAGCTATCCGCCCTGACTCGGTGGAACCCCTACAGACAGTAGAGACTGAT	3704		
OY	229	GlySerIleProAlaGluLeuArgGlnThrAlaLeuValAlaIleArgValValGlyTyrTyr	248		
Db	3705	GGAGACACCCCC-----	GTACCCGACCTT	3728	
OY	249	AspIleAspLeuSerIleIleArgAsnSerIleGlyValValAspValIleIleHisAsp	268		
Db	3729	CACGCACCAAGCACACACCCAGAAAGACGAGATGTGGAA-----	ATCATGAG	3778	
OY	269	LeuGlyGluProValTyrIleAspTyrArgAlaValaIleValArgGlyGlnGlyLysAsp	288		
Db	3777	GAGATGACGTC-----	GCATCTGTGACCAAGCAGGAGGGCACAGA	3818	
OY	289	AspLysAlaIlePheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis	308		
Db	3819	GCAGAGGAGCTTCTGCACAGAAAGAGGCCCTCAGACCTTCACGTTTGTGTCTCCAG	3878		
OY	309	HisGlyTyrGlnThrLysLysAsnLeuIleGluAsnAlaSerIleGlnHisGlyTyr	328		
Db	3879	GAA-----	GAACTAAAGAACATCAAAATGTGAAGACACTGTAGAGAT-----	3922	
OY	329	PheAspGlyArgTyrPleuAspArgSerValAspValIleLeuProAspAsnThrAlaAsp	348		
Db	3924	-----	ACAGATAAAGAGGTGTGCTGCTG-----	GAAACT	3950
OY	349	ValSerLeuIleTyrAspThr-----	GlyThrGln-----	TyrArg	360
Db	3951	GTATTCATTCTGTCAAGACTGAGGGGACTCAAGAGGCTGACCAGTAGCTGATGAGAA	4010		
OY	361	PheAspClnValValPhePhe-----	ThrIleAspProLysThrAsnGln	375	
Db	4011	ACCAAAAGACGTACCATTTTTCGAAAGCACTTGAAGGGTCTATATGAC-----	ACAGGC	4061	
OY	376	LeuThrThrAspProAspLysLeu-----	-----	383	
Db	4062	ATTAACAGTCAGTCGGAAAGGTCACGTGAAGTTGCCCTTAAAGGTGAAGGGACACAGAA	4121		
OY	383	-----	-----	383	
Db	4122	GCTGAATCTAAAAAGATGATGCTCTTGAACATGCGACAGTCACGCTAAGCTCTTCATCC	4181		
OY	384	ProValLysArgGluLeuLeuGlnGlnLeu-----	-----	Leu	394
Db	4182	CCCGTGGACAGAGATGGTGAAGTTCAAAGTCGAAAGGAGGAGAAACAGAACAGCCAAAC	4241		
OY	395	ThrValAsnMetGlyGlnAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu	414		
Db	4242	CATGTGAATGAAGAGAGCTTGAGCAACGAAACGCGTTACCGTA-----	TCTGACAG	4295	
OY	415	IleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPheProGluArgGlu--	433		
Db	4296	GTCAGTAACGACGCTCCACGACAGTGAATGTCGCCATCATGATGGGGCAAGAAAGTCC	4355		
OY	433	-----	-----	433	
Db	4356	AGCAGTTTGAGAGAGACCTCTCCCTCCCTAAGTCAAGAGGAGGAGTATGCACCAA	4415		
OY	434	---GlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThrGluPro	452		
Db	4416	ATTCAATTCAGAGCTGTAGAGGATCATTTACTCTTAACAGCGGCTGCAGAGAGGAAAG	4475		
OY	453	AlaClnValAspClnSerThrLeuGlnProValIleGlnThrValGlnLeuThrAspGly	472		

D	4476	GTCTTAGGCAAAACTGCCAATTTTAGAAACAGTGAACCTTGAGCCTTCAGGTCA	:	4535
O	473	ILEuMETAsPIeserProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu		492
D	4536	CATTTAGT-----CTGGAGAGAAATCTCTGGAAAAAATGAAAC-----		4577
O	493	AsnLeuValAlaAlaValArgHisLeuTyraSpmetProAspAspArgValLeuAla		512
D	4578	-----TTTGCCGCTCAT-----CCAGGGGAAGATGCTGGCC		4610
O	513	ILEasnhISAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer		532
D	4611	ACAGGCGCCGACTGTCAGGCAAAATGACACCGAGT-----ATACTATCT		4655
O	533	AlaValAlaAlaValAlaLeuProAspGlySerGluAsnGluValIleAspLeuProGlu		552
D	4656	GCTACTACCAAGAAAGCCTTAAGTCCGACCTGGAAAGGAGAC-----AAA		4700
O	553	ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyroGlnSer---LysLysVal		571
D	4701	ACCACATCTCAGAGTGAAGTGCATGAAAGTGCATGAGCAGAGTTCGTTGCCAGAGAGTGC		4760
O	572	ProLeuTyraValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyLyr		591
D	4761	AAAGTGAGTGAAGCAATTTAGGATTTAGAGCTGAAATAAGGATTTTGAACCTTGAGACC		4820
O	592	GlySer-----		593
D	4821	AAAAGCAGTAACTTGCACAAACATCATCCAGACAGCCGCTTGACCAGCTTTGTACGTCA		4880
O	594	---AspThrGlyThrArgLeuValThr-----LysPheGlnHisAsnLeuIleAsn		609
D	4881	GAAAGAAACAGCCACCGAAATGTTGAGCTGTGATTCACACACCAAGCTCAGCTGATAAA		4940
O	610	ArgAspGlyTyroGlnArgValArgLysLeuArgLysSerGluAspLysLysGlyValLys		629
D	4941	GCTACACAGCCAGACGCTGGACACGGAACGGAGAAAGAGAGAGAACCTCAGAGCCTCT		5000
O	630	LeuTyraIaThrLysProLeuSerHisProLeuAsnAspGlnLeuAlaGalaThrLeuGly		649
D	5001	GCACAGAGTGAACACCCCAATTACTTCAGCCAAAGAGAGCTCAGATCAACCGAGTGGGA		5060
O	650	TyrolGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeu-----SerThr		666
D	5061	CAA-----GCACATCTGTGATTTATTTCCAAAGACATGAGTGAAGACCTCAGAA		5105
O	667	ArgThrLeuGlnHisGluIleSerAspSerIleIleGlnAsn-GlyLysThrAsnArgThr		686
O	686	rTySer---LeuArgTyraLeuAspLysLysLeuTyroGlnAlaProGluThrTr		705
D	5166	GTCTCCCATCTGAGGAAGAGGAGGAGGTGAGCTGGA-ACAAAGTGTGGCCAGAA-----		5219
O	705	PglAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGly		725
D	5220	-----GAAAGATACAGAAAGTCACTAGTGAACCGAAGAGAT-----		5282
O	725	yValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyraGly		745
D	5244	-----GAAAGATACAGAAAGTCACTAGTGAACCGAAGAGAT-----		5282
O	745	naArgTySerLeuGlyValGlySerSerGlyLeuValSerAsp-----AlaAsnMe		762
D	5283	-----GAAAAAGGTGATGAT-----GTTGATGACCTCGAAAAACAGAACTC		5323
O	762	AlaIleAlaArgAlaGlyIleSerGly		771
D	5324	AGCCCTGGCTGATCTGATGATGCCCTCAGAGA		5351



Tue May 6 09:37:18 2003

us-09-914-168-2.rnpb

Page 65

Job time : 556 secs

---



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 5, 2003, 23:56:45 ; Search time 2019 Seconds

(without alignments)  
7371.794 Million cell updates/sec

Title: US-09-914-168-2  
Perfect score: 4/27  
Sequence: 1 MSKPVLFANRSMFVALAAY.....TGVEKGNPKLHFPIGTPF 919

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_LJ/USPRO\_spool/US09914168/runat\_28042003.151615.6799/app\_query.fasta\_1.1095  
-DB=EST -QFMT=fastap -SUFFIX=trst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bls0sum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR\_SCORE=Pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=Pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09914168 @CGN 1.1.1477 @runat\_28042003.151615.6799 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEDUQUEY -NEG\_SCORES=0 -WAIT -LONLOG -DEV\_TIMOUT=120  
-WARTIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST : \*  
1: em\_estba : \*  
2: em\_esthum : \*  
3: em\_estlin : \*  
4: em\_estlmu : \*  
5: em\_estlov : \*  
6: em\_estlpl : \*  
7: em\_estro : \*  
8: em\_hlcc : \*  
9: gb\_estcl : \*  
10: gb\_est2 : \*  
11: gb\_hlc : \*  
12: gb\_est3 : \*  
13: gb\_est4 : \*  
14: gb\_est5 : \*  
15: em\_estfun : \*  
16: em\_estom : \*  
17: gb\_gss : \*  
18: em\_gss\_hum : \*  
19: em\_gss\_inv : \*  
20: em\_gss\_pln : \*  
21: em\_gss\_vrt : \*  
22: em\_gss\_fun : \*  
23: em\_gss\_mam : \*  
24: em\_gss\_mus : \*  
25: em\_gss\_other : \*  
26: em\_gss\_pro : \*  
27: em\_gss\_rtd : \*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	153	3.2	792	14	B0506223	B0506223 EST613638
C 2	143	3.0	649	17	BH396320	BH396320 Ag-ND-161
C 3	129.5	2.7	620	12	BF942509	BF942509 1D2 CDNA
C 4	128.5	2.7	423	9	A1363592	A1363592 SMOVL2CAS
C 5	128.5	2.7	503	10	BE459389	BE459389 EST414681
C 6	127.5	2.7	3404	11	AK004785	AK004785 EST414681
C 7	124	2.6	2049	11	AT161466	AT161466 Homo sapi
C 8	121.5	2.6	494	10	AW191135	AW191135 68702D06
C 9	118.5	2.5	3897	17	BH770994	BH770994 LMGtag71
C 10	117.5	2.5	2582	11	BC011226	BC011226 Mus muscu
C 11	117.5	2.5	2607	11	BC032205	BC032205 Mus muscu
C 12	117.5	2.5	2837	11	BC032614	BC032614 Homo sapi
C 13	116	2.5	607	10	AV557254	AV557254 AV557254
C 14	113.5	2.4	503	11	AT108685	AT108685 Zea mays
C 15	112	2.4	544	9	A1519730	A1519730 LD39445.5
C 16	111	2.3	3144	11	BC025329	BC025329 Homo sapi
C 17	110.5	2.3	678	12	BF491509	BF491509 AT28265.5
C 18	110	2.3	1673	11	BC013085	BC013085 Homo sapi
C 19	109	2.3	1501	11	AT104509	AT104509 Zea mays
C 20	107	2.3	2083	11	BC022490	BC022490 Homo sapi
C 21	106	2.2	588	10	AW053211	AW053211 614073B08
C 22	105	2.2	758	12	BE786830	BE786830 601475986
C 23	105	2.2	1557	107	BH770729	BH770729 LMGtag47
C 24	104.5	2.2	1021	17	CNS071R5	AL425271 clone BAO
C 25	102.5	2.2	1364	10	AW342205	AW342205 GtHST91
C 26	102.5	2.2	1867	11	BC019335	BC019335 Homo sapi
C 27	102.5	2.2	1876	11	BC018883	BC018883 Homo sapi
C 28	102	2.2	1014	13	B1086482	B1086482 602849752
C 29	102	2.2	2039	11	AK013402	AK013402 Mus muscu
C 30	101.5	2.1	513	17	AZ927049	AZ927049 476.d1s16
C 31	101	2.1	2340	11	BC028159	BC028159 Homo sapi
C 32	101	2.1	2936	11	AK011769	AK011769 Mus muscu
C 33	100.5	2.1	912	9	AL546961	AL546961 AL546961
C 34	100	2.1	800	12	BF256095	BF256095 HVSMEF000
C 35	100	2.1	1041	9	AL523772	AL523772 AL523772
C 36	100	2.1	3499	11	AK019509	AK019509 Mus muscu
C 37	99.5	2.1	969	17	CNS013XL	AL103443 Drosophi
C 38	99.5	2.1	1514	11	AT108441	AT108441 Zea mays
C 39	99.5	2.1	1562	11	AK002434	AK002434 Mus muscu
C 40	99	2.1	1022	14	W29713	W29713 mc05a09.f1
C 41	99	2.1	1970	11	AK010097	AK010097 Mus muscu
C 42	99	2.1	2345	11	AK014398	AK014398 Mus muscu
C 43	99	2.1	4379	11	BC008524	BC008524 Mus muscu
C 44	98.5	2.1	884	17	BH137735	BH137735 ENTNI81TF
C 45	98.5	2.1	1409	11	AT109022	AT109022 Zea mays

## ALIGNMENTS

RESULT 1  
B0506223/c  
LOCUS B0506223 792 bp mRNA linear EST 22-JUL-2002  
DEFINITION EST613638 Generation of a set of potato CDNA clones for microarray  
analyses mixed potato tissues Solanum tuberosum CDNA clone STMG192  
3' end, mRNA sequence.  
ACCESSION B0506223 GI:21922105  
VERSION B0506223.2  
KEYWORDS EST.  
SOURCE potato.  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 792)







Ox	897	AspValAlaThrGlyValIleuLysIleuLysIleuProIleuLeuNH <sub>2</sub> Phe	913
Dd	462	GAGTATGCCCTTCATGCACCGCACCGGG-----AGGTTCACTTC	503
RESULT 6			
LOCUS	AK004785		
DEFINITION	AK004785 Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:120015F06; myosin phosphatase, target subunit 1, full insert sequence.	3404 bp mRNA linear HTC 19-JAN-2002	
ACCESSION	AK004785		
VERSION	AK004785.1 GI:12836227		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male lung cDNA to mRNA, clone:120015F06.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Mech. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitasuna, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuru, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, T., Izawa, M., Nishi, K., Kiyosawa, H., Kondou, S., Yamanaoka, I., Salto, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Salto, R., Kadote, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiya, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, T., Pesole, G., Quackenbush, J., Schriml, L. M., Stabill, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., But, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Humé, D. A., Kamuya, M., Lee, N. H., Lyons, P., Matchanlon, L., Mashimo, J., Mazzarelli, J., Mommaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyokawa, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kontsuki, S. and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
MEDLINE	21085660		
PUBMED	11217851		
REFERENCE	5 (bases 1 to 3404)		
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,		

TITLE	JOURNAL	COMMENT
<p>Submitted (10-JUL-2000) Yoshinide Hayashizaki. The Institute of Physiol and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Shinohe-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)</p> <p>Please visit our web site (<a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a>) for further details.</p> <p>cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGCGCCGACACGACGATGATTTTATTTTATTTVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 5' GGAGAGACAGAGATCCAGAGCTCAATTAATTAATTAACCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.</p>	<p>Location/Qualifiers</p> <p>1. 3404</p> <p>/organism="Mus musculus"</p> <p>/strain="C57BL/6J"</p> <p>/db_xref="FANTOM_DB:1200015F06"</p> <p>/db_xref="MGI:MGI:1904368"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="1200015F06"</p> <p>/sex="male"</p> <p>/tissue_type="lung"</p> <p>/clone_lib="RIKEN full-length enriched mouse cDNA library"</p> <p>/dev_stage="adult"</p> <p>1. 3404</p> <p>/gene="Mypt1"</p> <p>1. 3015</p> <p>/gene="Mypt1"</p> <p>/note="data source:MGI, source key:MGI:1309528, evidence:ISS</p> <p>myosin phosphatase, target subunit 1 putative"</p> <p>/cocon_start=1</p> <p>/protein_id="BAB23563.1"</p> <p>/db_xref="GI:12836228"</p> <p>/db_xref="MGI:MGI:1309528"</p> <p>/translation="MKMAADAKOKRNDPDKRMISGFTDLEPPYVKRQTKYKFDGAVFLAACSSGDTDEVLKLRGADINRANVDGLALHQAQCIDDNVDMVKELVANGANINP DNEGVIDEARKAEEERVMIPARQWLNSGHSIDYVSHAASGGLAHVAHKAATYEVLLKIL IOAGYDNIKDYDQWPTPLHAAAHMGKEEACRILVDNLCDEMTVAKVCGOATFADVEDDI LGYEELEOKKOTDILHSEKRPDKSPILSTAMENNOPOKAFKNKETLILIEPKNASRIT ESLEHKADEEBEERKDESSSSSEDEDEDESEAEFDKTRPMASVSNATHSQOAP AAATAPLSSNQGTPTSPVKRFPISSTIKISKKEERKDESPASWRDLKRTGSGALA EISASKAEOKEDTGAVMKRSASPRLSLSDNKKREKDMNGTRLAAYTPPIPRLASTI SDIEENKRESSSLRTSSSYTRKMWEDLKNSSINGSTYHRSQGRQDILISCS VPSTPTPTYSAGLQRLSPSSSTAKPPGSSAGTOSSTNRLAAEDSTKEKD SAPATVIVPAAPVYVNAAPSTTLTITTTGTVSEVERRRSYLTPVDEESSOKA RSROAROSRSTGCVTLTDIOEAKTIGRSKSTTRQDENREKKEKKEKODKOE KKEEASREDEYKORYSRTYDETTRRNPVSTSSSSSSSSSLSLSTLYASSQLNR</p>	<p>FEATURES</p> <p>SOURCE</p> <p>gene</p> <p>CDS</p>







```

Db 838 CATAGGACGAGTGTGTGATGATCCACAGTGGAGCAGAGGCTTGGCCACCAAGTA 897
Oy 456 AspluSerThrluGluproValIleGlulThrlValGlulThrlAspGlyIleLeuMet 475
Db 898 GCCACACATGCGCTG---GTAAGCATATGAGACAGCCATGGAAGAGACAAAGATTATAGTC 954
Oy 476 AspluSerProIleGlupheSer-----AlaSerAsnleuIleGlnAspLysLeuAsn 493
Db 955 AATGATCGCAGTGTGCTTGTCTCTGATGCTTCCCGAGAGGTTCAGAGACTATCTGAG 1014
Oy 494 LeuValAlaLysAlaAlaGlnHlsLeuTyraSmetProAspAspAlaValAlaIle 513
Db 1015 GGAATGGCAGCTGCTGGCAACTATGCTGG-----1044
Oy 514 AsnHlsAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAla 533
Db 1045 -----GTCACACCGCTTTCATGACCTTTTACCCGCTCAGGCTTTTCGCC 1089
Oy 534 ValAlaArgAlaIleLeuProAspGlu-----542
Db 1090 AAGCTTTCACACACACACACCCGCTTCCCT-----CCTCAGCATCCCTCATTTGCT 1149
Oy 543 -----SerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558
Db 1150 AACATTGCCAAAGTGGAGCAGCATGTGTGGAGAAAGAGACGAGCATGTTAGTACAC 1209
Oy 559 ArgLysThrProAlaAspValTyraGlnSerLysLysValProLeuTyraIlePheValAla 578
Db 1210 AGCAAGGATCCACCCGCTTCCCT-----CCTCAGCATCCCTCATTTGCT 1257
Oy 579 SerAspLysProArgAspGlyGln---IleGlyLeuGlyLysThrAlaLeuAlaIle 597
Db 1258 GTTGATTACCACTGACTGACAGCAGCATGCTCTATTGCTGACCATGGAGACCTGTAGT 1317
Oy 598 ArgLeuValThrLysPheGlnHlsAsnLeuIleAsnArgAspGly-----TyraGln 614
Db 1318 TATGCTTCTTACTGCGACTGACAGGAGCATGACGACCTTGGAGACACCTGTCAATGA 1377
Oy 615 AluGlyAlaGluLeuArgLeuSerGluAspLysGlyVal 628
Db 1378 GCGGCGCTGCTATTCGCCGAGCAAAATCTGACGTAATTTA 1419

RESULT 8
AM191135 494 bp mRNA linear EST 30-MAR-2000
LOCUS 687022D06.x1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
DEFINITION sequence.
ACCESSION AM191135
VERSION AM191135.1 GI:6465676
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 494)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNPUBLISHED
CONTACT: Walbot V
DEPARTMENT: Department of Biological Sciences
STANFORD UNIVERSITY
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 687022 row: D column: 06.
FEATURES
source
1..494
/organism="Zea mays"
/cultivar="Illinois High Oil"
/db_xref="taxon:4577"

```

```

/clone.lib="687 - Early embryo from Delaware"
/tissue_type="embryo"
/dev_stage="14, 21, 28, and 35 days after pollination"
/lab_host="E. coli SOLR"
/notes="Organ: embryo; Vector: pBluescript SK; Site: 1: XhoI
; site 2: EcoRI; Library was prepared by Stagen using
the Uni-ZAP XR system (Stratagene BN937328-12) Clones
were picked by a Q-dot after blue/white selection
(amplified in resistance - use 100 micrograms/microliter).
Developed from a pool of equal amounts of RNA from
developing embryos sampled at 14, 21, 28 and 35 days after
pollination of the Illinois High Oil maize Strain Cycle
90. This closed strain has been selected for high oil
concentration for 90 generations and originates from the
1890s era open pollinated variety Burr's White"

BASE COUNT 147 a 147 c 97 g 103 t
ORIGIN

Alignment Scores:
Pred. No.: 0.000932 Length: 494
Score: 121.50 Matches: 41
Percent Similarity: 42.458 Conservative: 18
Best Local Similarity: 29.508 Mismatches: 49
Query Match: 2.57% Indels: 31
DB: 10 Gaps: 6

US-09-914-168-2 (1-919) x AM191135 (1-494)

Oy 789 ThrGlyGlyIleGlnAlaGlyTyrIleTyrPserAspAsnPheAsnHlsValProTyraArg 808
Db 448 TCTGAGGCTCACATGAGGGA-----AATTT-----CCACCTCAT 413
Oy 809 LeuArgPhePheAlaGlyLysArgInsleIleArgGlyTyraIleHlsAspSerLeuSer 828
Db 412 GAGCATTGTCGCAATGTGTGGACAAATAGCTAAGAGATAT-----371
Oy 829 ProIleSerAspLysGlyTyrIleThrGlyGlyGlnValLeuAlaValGlyThrAlaGln 848
Db 370 -----GAAAGAGGTGCTGTGGCTGCTGCTGCTTCTTATCTGTAGTGTAGTGA 320
Oy 849 TyraSerTyraGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAla 868
Db 319 GTGTCAATGCCGCAATGTTTGGGCCCCCTGGAGAGCGCTGCTGTGGAGCATGATGTAGT 260
Oy 869 TyraSpLysGly-----PheThrAsnAsp-----Thr 877
Db 259 CTTGCTTGTGCTGCTAAGTCTCTGTGACCCAGCGAGCTGTGGAAGCCAGGAGCT 200
Oy 878 LysIleGlyAlaGlyValGlyValArgTyrPalasPheProValGlyGlnValArgValAsp 897
Db 199 GGGTTTGCTATGAGTGTGGCGTCCCGCTGGAGCTCTGAGACCTTACGCGCTTGAG 140
Oy 898 ValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHlsPheHeIleGly 916
Db 139 TACGCTTTAATGATAGGCAAGCAAGC-----AGATTTCACCTTGTGTGTGGC 92

RESULT 9
BH770994 3897 bp DNA linear GSS 01-MAY-2002
LOCUS BH770994
DEFINITION LMGtag718 MG1363 Random Sequence Tag library Lactococcus lactis
subsp. cremoris genomic, DNA sequence.
ACCESSION BH770994
VERSION BH770994.1 GI:20373951
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris.
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 3897)
AUTHORS Bolotin, A., Ehrlich, S.D. and Sorokin, A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Aliments, (2002) In press
CONTACT: Sorokin A

```

Genetique Microbiome  
 INRA  
 CRJ INRA, Domaine de Villvert, 78352 Jouy en Josas cedex, France  
 Tel: 33 1 34 65 25 16  
 Fax: 33 1 34 65 25 21  
 Email: sorokine@jouy.inra.fr  
 best homologue in strain IL1403 is dneE (92%)  
 Class: shotgun  
 High quality sequence start: 30  
 High quality sequence stop: 3867.  
 Location/Qualifiers  
 1. 3897  
 /organism="Lactococcus lactis subsp. cremoris"  
 /strain="MGI363"  
 /db\_xref="taxon:1359"  
 /clone\_lib="MGI363 Random Sequence Tag Library"  
 /note="Vector: pSGM2; Site\_1: SmaI; Library of  
 chromosomal fragments of L.lactis strain MGI363 was  
 prepared by partial AluI digestion or by sonication."  
 BASE COUNT 1211 a 628 c 848 g 1210 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0834 Length: 3897  
 Score: 118.50 Matches: 122  
 Percent Similarity: 34.35% Conservative: 114  
 Best local Similarity: 17.76% Mismatches: 236  
 Query Match: 2.51% Indels: 215  
 DB: 17 Gaps: 32

US-09-914-168-2 (1-919) x BH770994 (1-3897)

QY 199 IleglySerSerHiglnylThrGluProTyrrAlaAsnIleLysAlaLeuGluAsp 218  
 ||||| : : : : : |||||  
 Db 1680 ATTGGTTTACAGAGAAGAGAGAGAAATGTTT-----GCACCTTAATACA 1727  
 QY 219 IlethrgIngluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgInthra 238  
 ||||| : : : : : |||||  
 Db 1728 AAACACAGATACAGTCTTTTAAAT-----ACTGTTGTTAACTA---CAGCATTA 1775  
 QY 239 LeuValAlaAlaArgAlaValAlaGlyTyrrAspIleAspLeuSerIleLeuArgAsn 257  
 ||||| : : : : : |||||  
 Db 1776 TTGGAGAGGCGCTTACGTTAGCGTATAAACGTCGGAATTTGATGGGAAACCTC 1835  
 QY 258 -----SerIleGlyLysValAspValIleIle 266  
 ||||| : : : : : |||||  
 Db 1836 CATGACGCTTTGCGTTGTCTCAGAAAGCTCAATGCTTAATTACACACAGCTTTTACT 1895  
 QY 267 HisAspLeuGlyGluProValTyrrIleAspTyrrAlaValAlaGluValArgGlyGlu 286  
 ||||| : : : : : |||||  
 Db 1896 ATTGCATTAATAAT-----TTTGAAATGCGCGGCTTCCAGTT----- 1931  
 QY 287 AlaAspAspLysAlaPheThrThyValAlaAspGluValProLeuLeuIleGlyAspVal 306  
 ||||| : : : : : |||||  
 Db 1932 -----TCCTTTTCTTTTATTTGCTTAATAATACA----- 1958  
 QY 307 PheHisIleGlyLysTyrrGluThrLysAsnLeuIleGluAsnAlaSerAlaGluHis 326  
 ||||| : : : : : |||||  
 Db 1959 -----GAAAGTTATATAAAACTTCTT-----CGCATTTCAACATTTTCA 1997  
 QY 327 GlyTyrrPheAspGlyArgTyrrLeuAsp-----ArgSerValAspValIle 341  
 ||||| : : : : : |||||  
 Db 1998 AATTATGATGACGCGCATTTTCTGATATTCATAATCATTTTGTCCAGATCCACTGGT 2057  
 QY 342 LeuProAspAsnThrAlaAspValSerIleuIleTyrrAspThrGlyThlGlnTyrrPhe 361  
 ||||| : : : : : |||||  
 Db 2058 ATTCCAGAACTTATGGACTTTAGTGAGCTATCAGACTTCAGCA-----TTT 2108  
 QY 362 AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThyAspProAsp 381  
 ||||| : : : : : |||||  
 Db 2109 GCCCAGCAGAGCTATATTGGATTACACAGACAGATATA----- 2150  
 QY 382 LysLeuProValLysArgGluLeuLeuGluInLeuLeuThrValAsnMetGlyGluAla 401

Db 2151 -----GGAACCTAAA 2159  
 QY 402 TyrrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrrPheAsn 421  
 ||||| : : : : : |||||  
 Db 2160 TTTAATTATCAACACCTT-----CCTTTCCGACGTTGCTTATCTGAAT 2204  
 QY 422 MetValAsnThrGluIleValPheProGluArgGluInIleGlnAsnAspGlnValSer 441  
 ||||| : : : : : |||||  
 Db 2205 TTTGCTGATATAGAAATTTTA----- 2225  
 QY 442 PheGluInSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461  
 ||||| : : : : : |||||  
 Db 2225 ----- 2225  
 QY 462 ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGlu 481  
 ||||| : : : : : |||||  
 Db 2236 GCTATTTTGCACGCGCTGCT-----GATGGGCTTCTTCTTGACGAAAGT---TTGACG 2276  
 QY 482 PheSerAlaSerAsnLeuIleGln----- 489  
 ||||| : : : : : |||||  
 Db 2277 GTCAGTTCAATGAATTTGCTCCAAAGCCACAGACTTATGAATCTTATTTCCAGAAATAT 2336  
 QY 490 -----AspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrrAsp 504  
 ||||| : : : : : |||||  
 Db 2337 TTTCGCAAGCAGCTGCAAAATCTGTCAGTATTAACAGCAAAATTCATTATCAATTT--- 2393  
 QY 505 MetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeu 524  
 ||||| : : : : : |||||  
 Db 2394 -----GATGAAATAATGTGACCTCCAGACTTTCATTAACAAAGCAACACCTTCAA 2444  
 QY 525 GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu 544  
 ||||| : : : : : |||||  
 Db 2445 CATCTCCAGAGAAGACCATTTTAGACTGTCTGCTCCACTTGAGAAAGAAAAAATCACTG 2504  
 QY 545 AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAsp 564  
 ||||| : : : : : |||||  
 Db 2505 ACAGAAATTTGCTGCTAGTAAGTCTGCAAGTGTCTTAGCGTAGAAGATCAACAGCATGAT 2564  
 QY 565 -----ValTyrrGln-----SerLysValProLeu----- 573  
 ||||| : : : : : |||||  
 Db 2565 TTGTCAGTATATACAGACCGCTTACGAAAAAGAGTTTGATGATCCACCAATGGGATTT 2624  
 QY 574 -----TyrrValPheValAlaSerAsp-----LysProArgAspGlyGln 586  
 ||||| : : : : : |||||  
 Db 2625 GATGATATTTTATTTATTTGCTGCGGATTTTGTGCCCTATGCCCGTGAAACACATTTAT 2684  
 QY 587 IleglyLeuGlyTyrrGlySerAspThrGlyThrArgLeu----- 599  
 ||||| : : : : : |||||  
 Db 2685 TCGGGAATGGGCGTGGGCTGCTGCGCTTGTGGCTATGTTGGGAATCACT 2744  
 QY 600 -----ValThrLysPheGluHisAsnLeu-----IleAsnArgAspGlyTyrr 613  
 ||||| : : : : : |||||  
 Db 2745 CAGGTTGACCCCTGCAACATAATCTTTCTTGAACGTTTTTTGAACCCACAAAGAGTNG 2804  
 QY 614 ---GlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyrrAla 632  
 ||||| : : : : : |||||  
 Db 2805 GCCATGCCCGATATTGATATGATATGATGCCCGATGACCGTGGCTGAGTGTGGCCTAT 2864  
 QY 633 ThrLys-----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGly 649  
 ||||| : : : : : |||||  
 Db 2865 ATGAAATAATCGTTATGTTACAGACCATGTAGACAAATGTACATTTTCAACATTTAGCA 2924  
 QY 650 TyrrGlnGln-----GluValPheGlyHisSerThrAsnGlyPheAsp 663  
 ||||| : : : : : |||||  
 Db 2925 AAAAGACAGGCGCTCAGAGATGTCGTAAGCCTTTGGGATGACCGAAGCTGAATTAATCT 2984  
 QY 664 LeuSerThrArgThrLeu-----GluHisGlu----- 672  
 ||||| : : : : : |||||  
 Db 2985 AATTGACTCGGATGCTTGTCCGCCGTTTGGAACTGTGCGCTGATGAATATGAATAATAT 3044  
 QY 673 -----IleSerArgSerIleIleGlnAsnGlyGlyTyrrAsnArgThrTyrrSerLeuArg 690  
 ||||| : : : : : |||||

```
Dbb      3045 CACAGCTTTTAAAGCCGAAGATATTGAAAAATGTAAAATTAACAAGATTATGAGACTCGCC 3104
Oy       691 TTTArGLeAsPvLyLeuLysThrGlnAlaProProGluThr-----Trp 705
          |||:::~::~|||
Db      3105 AGACGAATTTGAGGGAATGGCTCGCCAGACTTCACACATGCCTCAGSGTGTTGTTCTTCT 3164
Oy       706 GlnAsPleuProValAsPheValAsnGlyLysProSerGlnIuaLaLeuAlaGly 725
          ::::~:::::|||||:::~::|||
Db      3165 GAAGAATCTTTTGATCAATAATTATGTCATTAATAAACCGTCAGAAGATTAGCTTGCACAA 3224
Oy       726 ValAlaValHisLysThrValAlaLaspAsnLeu----- 736
          ||||~|||
Db      3225 TATGAGCACCAAGATGTGGAACGATAGCGCTCTTA AAAATTGACTCTCGGATTAAGT 3284
Oy       737 -----ValAsnPromeTarGcLYr-----ArgInArGtyrSerLeu 749
          ::::~:::::|||:::~::|||
Db      3285 AATTTAACGATCATTTTCGAGATTACGTCGATTGTCGTA AAAAACGTCACAAAAATTGATATT 3344
Oy       750 GluValIGlySerSerGlyLeuValSerAspAlasnmEtAlaIleAlaArgAlaGly--- 768
          ::::~:::::|||||:::~::|||
Db      3345 GACCCTTAAAAAATAGATTGAAAGATGAACAACGACATTGGCACCTTTTAGAGCTGGAAT 3404
Oy       769 IleSerGlyValTyrrSerPhe 775
          |||~:::~::|||
Db      3405 ACGATGGGAATTTTTCAAGTTT 3425
```

LOCUS	BC011226	2582 bp	mRNA	linear	HTC 07-AUG-2002
DEFINITION	Mus musculus, Similar to KIAA0244 protein, clone IMAGE:418839,				

ACCESSION	BC011226
VERSION	BC011226.1
	GI:15029980

SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE  
1 (bases 1 to 2582)  
Mammalia; Eutheria; Rodentia; Scleroglossa; Muridae; Mus

Authors	Straussberg, R.
Title	Direct Submission
Journal	Submitted (25-JUL)

USA  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
gene collection (MGC), Cancer Genomics Office, National Cancer

REMARK	COMMENT
NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>	
Contact: MGC help desk	
Email: <a href="mailto:coarbs-r@mail.nih.gov">coarbs-r@mail.nih.gov</a>	

Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Assayed by: Robert M. Green, M.D.

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Web site: [http://www.  
Contact: amgebcm.tmc](http://www.Contact: amgebcm.tmc)

Yoon, V.S., Kowis, C.R., Lawrence, S., Marlin, R.G., Muzny, D.M.  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.inl.gov>

This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene

**FEATURES**  
This clone has the following problem: Incomplete processing  
location/Qualifiers

FEATURES	Location/Qualifiers
source	1. .2582

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="IMAGE:4188839"

```

```

/tissue_type="salivary gland, 10 week old female mouse"
/clone_id="NCI_CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT      883 a      482 c      620 g      597 t
ORIGIN

```

Alignment scores:	
Pred. No.:	0.0531
Score:	117.50
Percent Similarity:	34.218
best Local Similarity:	18.83%
Query Match:	2.49%
DB:	1.1
Gaps:	4.1
length:	258
Matches:	158
Conservative:	129
Mismatches:	262
Indels:	291

US-09-914-168-2 (1-919) x BC011226 (1-2582)

```

Oy 108 AspmcsevalIlegIugIuThrThrProIeuSerIeugIuIeuPhealagInIu 127
    ||| ||| :: |||||::: :: :: ::

```

QY 128 SerThrGluMetGlyIleAsnProAsnAspTyr--IleProGluTyrGlnGluGluGln 146

Db 178 GATCCTATGCTAGGATCTGCAGTAACCACTTCTGTTTGCTGTTTGGATAGCATGAT 237

	17	111111		Seigunval	131
	27	111111			
Db	238	CCCAATTTCAGATGCGCTGTTCACACAGTGTGGTCTTGACGATATTATGATGAAGA	297		

QY 152 ValValPro-----ProThrLeuGluProGluLysProGluLysLeuLysArg 1677

```

DB      GGTGTTAAAGAAAGTCGTAAATGATACCAATTCATGATGAGAGCAAGCAATTGATTTCCTAACACAGG  357
OY      168 ---[LeuTYR]AARDLeuPheAsnAspG]VVA]-----AsnIsvValProARLeu 1833

```

Db 358 AGTTTGAGGGACAGAGAGTAGAAGACAATTCAGTAGATCACCAGAAGAAATGCACCTCGTTTA 417

Dh  
A18 ATCCGCGACAGCAGCAACGATACCTGTTTGGCGAACCGGCTAAATTCGA 477

QY 184 LysAlaLys-----PheTryGI nSerSer----- 191

QY 192 -----GlnSerArgGlu-----ThrSerAlaIleGly--- 200

Db 478 GCACTCTAAGCACCACAAAAGCCATCTGGGAGACCTTATCTACCTCTAAAGTAGGGTG 537

Db 538 AAGCCACGAAGATGTCAGGGTAAAGAGAAGTTTATGCCCTCAGTGAATCTGAACAC 597

Qy	214	-----AlaAlaLeuGluAspIleThrGlnGlu---	222
		:	

223 Ser1ametaspleungilyseriIleproArq-----LeuArgGlnThralaleu 239

Db 658 .TCAGCATCTTCAGTTCATCTTCAGTTCGCTGTGTCGGAATGAAGAGAGGCT--- 714

Db 715 -----GAGTTTGATCCCAACATGATGCAATGCAATATATACG 742  
 09 240 ValaAaAaAaAaGAlaGAlGyIyIyIyIyAspIleAspLeuSerIleIleIleArgAsnSerIle 259

260 GlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAla 279

Db 748 GGTAAGTGAACGTGCATCTCCTGAGTTAGATTGTCCACTTCTCTCAGAGACTAGTGCT 807

Db 808 AGTGTGAGAGAGAACATTGAGGCTCTGATGTGAATTAAGCTAAGACT----- 858

QY	297	AspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys	316
:::			::: ::: :::

```

Oy 317 AsnLeuIleGlu-----AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrp 333
Db 907 GATCTTGTTCGGGAGTACTGACACCATTTGAAGAAAGATGACGGGATAAAGCT 966
Oy 334 LeuAspArgSerValAspValIleLeuPro-----AspAsnThrAlaAspValSer 350
Db 967 GATCAAGATTCAGAGAGGTGAAGTTTCCCTCGAAGTACAGCAGCGCCGAG----- 1020
Oy 351 LeuIleTyrAspThrClyThrGlnTyrArgPheAspGluValValPheIleThrLeuAsp 370
Db 1021 -----GAA 1023
Oy 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuAsn 390
Db 1024 CCCGATCTCTTGATGTGCTAGT-----GACTCTGCTTGGCCCATATAAATAAGAGA 1077
Oy 391 GlnGlnLeuLeuThrValAlaSerMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeu 410
Db 1078 GAAAAG-----AATGAAAGGGCAGAAATGTCATTGTGGA-----TTG 1113
Oy 411 SerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrClyIleValPhePro 430
Db 1114 AAGAAATACCTG-----GATATTGTGATTAACCTGAGAACTGACCT 1155
Oy 431 GluArg-----GluGlnIleGlnAsnAspGlnValSer 441
Db 1156 CAGAGAAATGAGTGGGAAACACTGGGCTATGGTGGAGCAGACAGGCTAATGATGCTCAG 1215
Oy 442 PheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461
Db 1216 TTACAGAGCAGCAGAGTTCATAGTCAGATTGGAGGAAGTGCATTCCTGCTTTGAA 1275
Oy 462 ProValIleGlnThrValGlnLeuThrAspGlyIleLeuMetAsp----- 476
Db 1276 CCGGAAGCTACGACCTTAAGAAATACTATTGTGGATGCTTTCACCAAAATTCAAAGCAA 1335
Oy 477 -----IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn 493
Db 1336 TTGAATATTACTCAAGATTAATGAAGCAACCACTTCAGAGTATAGTAAGTGGT 1395
Oy 494 LeuValAla-----AlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
Db 1396 TTAGAGCCCCAAAACATAAAACCCCAACATAAATCTGTAATCATCTTAACAACAAAGC 1455
Oy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArg 521
Db 1456 ATGACCACAGACAGACCCCAAAAACTGTGGCAGCAAGCATGAAGTACTATACAAA 1515
Oy 522 SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAsp 541
Db 1516 ACT-----AAATCTAATGTCAAGCGCTGTCAAGCGA-----AAT 1548
Oy 542 GluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
Db 1549 TCTGATGACCCGAGACCCAGCCCTGATTTCTCAGAGGCCAGTCAAACTGAGAAAAACAA 1608
Oy 562 ProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLys 581
Db 1609 GGTGATTAAGTTTGAAGAGTCA----- 1632
Oy 582 ProAlaAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 1633 -----AGTTGTAATTCGAGCTTAATCTGTGAAA 1662
Oy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuAspLeu 621
Db 1663 ACCCAAGCTCATTTCTGTATTGAAA-----ACAATG 1692
Oy 622 SerGluAspLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1693 CCCCAAGATCAAAACACATGCAAGATT-----TCCAACCTTAATCATCTCATAGT 1746
Oy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661

```

```

Db 1747 GATAAGCTTCAT-----GATCAGCT----- 1767
Oy 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 1767 ----- 1767
Oy 682 GlyTyrAsn-ArgThrTyrSerLeuArgThrArgLeuAspLysLeuLysThrGlnAlaPr 701
Db 1768 GGCTTCTTAAGAGGCTCCCTCCATCTGTACAACTGGAC----- 1807
Oy 701 oProGluThrTyrGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAl 721
Db 1808 -----ACTTGG-----TGCAATTCAGTCAAAAA-CAGAGCCAGAAACCTCAGACAGCGC 1856
Oy 721 AlaLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet-- 740
Db 1857 AGCTTGCGGGAAGGGCAGTACCAT-----GTGAAGGATGAGCATGACCATTCAGTCA 1910
Oy 740 ----- 740
Db 1911 TGAGCATTTTAAGAGAGCATTAACCTGAAGCCAGAAAAACCTGATAGCACTGCAGCC 1970
Oy 741 -ArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAl 760
Db 1971 TCGCCAAAGAGAGAGCAGCCGAAGTTTCTGTGATGAGCCCTCATGTGTCAATGCCAGA 2030
Oy 760 AsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 780
Db 2031 CAACATAGTACTGTAAA----- 2049
Oy 780 rGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrPheAs 800
Db 2050 -----AAGAAGGCTCAGATCAGACACCACTTAATGAAGCAATATATATGGCACTGC 2102
Oy 800 pAsn-----PheAsnHisValProTyrArgLeuArgPheAsnAlaGly 814
Db 2103 AAGCAAGCAATGTGCGTTTGCMAAAACACATGCAACAGGTATATGTCGCGC 2157

```

RESULT 11  
 BC032205 2607 bp mRNA linear HTC 07-AUG-2002  
 LOCUS BC032205 Mus musculus, clone IMAGE:5342620, mRNA.  
 DEFINITION Mus musculus, clone IMAGE:5342620, mRNA.  
 ACCESSION BC032205  
 VERSION BC032205.1 GI:21618986  
 KEYWORDS HTC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 2607)  
 TITLE Direct Submission  
 AUTHORS Strausberg, R.  
 JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov  
 CONTACT Contact: MGC help desk  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Galthersburg, Maryland;  
 Web site: http://www.nisc.nih.gov/  
 Contact: nisc-mgc@nhgri.nih.gov  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
 Maduro, Q.T., Mastello, C., Masker, B., Mastrian, S.D., McCloskey, J.C.,



```

QY 562 ProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLys 581
    |||:|||||:|||||:
Db 1628 GGTGATAAGTTTGGAGAAGTCAAG----- 1651
QY 582 ProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThr 601
    |||:|||||:|||||:
Db 1652 -----AGTTGTAATTCGTGAGTTAAATCTGTGAAA 1681
QY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuValGlu 621
    |||:|||||:|||||:
Db 1682 AGCCAAAGCTCATTCGTATTGAAA-----AGAAATG 1711
QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
    |||:|||||:|||||:
Db 1712 CCCACAGATCAAAACACATGACAGATT-----TCCAAACCTTAACTCATCTCATAGT 1765
QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlnHisSerThrAsnGly 661
    |||:|||||:|||||:
Db 1766 GATAAGCTTCAT-----GGTCACTCT----- 1786
QY 662 PheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGlnAsnGly 681
    |||:|||||:|||||:
Db 1786 ----- 1786
QY 682 GlyTyrAsn-ArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaPr 701
    |||:|||||:|||||:
Db 1787 GGTCTCTCAAGAGAGCTCCCATCCGTACAAACTGAC----- 1826
QY 701 oProGluThrTyrGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAl 721
    |||:|||||:|||||:
Db 1827 -----ACTTGG---TGCAATTCACAGTCAAAA--CAGAGCCAGAGACCTCAGCAGAGCC 1875
QY 721 AluLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet-- 740
    |||:|||||:|||||:
Db 1876 AGCTGTGGAGAGGACAGTACCCAT-----GTGAAGGATGAGCATGACCATCCAGTCAAG 1929
QY 740 ----- 740
Db 1930 TGAGCATCTTAAGAGAGACGATAAAGTGAAGCCAAAGAAACCTGATAGAACCTGCAGCC 1989
QY 741 -ArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAl 760
    |||:|||||:|||||:
Db 1990 TTCCTCAAGAGAGACGCCAGAGTTTCTCTGGATGAGCCATCCATGTTTCATCCAGCA 2049
QY 760 AsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTy 780
    |||:|||||:|||||:
Db 2050 CAACATAGCTACTGTATAA----- 2068
QY 780 rGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAs 800
    |||:|||||:|||||:
Db 2069 -----AAGAAAGGCTCAGATCAGACCCCTCAATTGAAGAAACAAATATATGTGACTCC 2121
QY 800 PAsn-----PheAsnHisValProTyrArgLeuArgPhePheAlaGly 814
    |||:|||||:|||||:
Db 2122 AAGCAAGCAATGTGGGTTTTCAGAAAAACACATGAGCAAGGTTTATGTGTGCGC 2176
RESULT 12
LOCUS BC032614 2837 bp mRNA linear HTC 26-JUN-2002
DEFINITION Homo sapiens, similar to RAB5 interacting protein 3, clone
IMAGE:5557263, mRNA.
ACCESSION BC032614
VERSION BC032614.1 GI:21595728
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2837)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian

```

```

REMARK
COMMENT
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: c9abbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCPD/DPF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@hgrl.nih.gov/
Akhter, N., Ayele, K., Beckstrom-Stenberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granice, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R.,
Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantieri, S., Thomas, P.J., Touchman, J.W.,
Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAC Plate: 69 Row: m Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gl: 13376344
This clone has the following problem: frame shifted.
FEATURES
source
1. 2837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557263"
/tissue_type="Skin, melanotic melanoma."
/clone_id="NIH-MGC_72"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
BASE COUNT 581 a 1000 c 805 g 451 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0627 Length: 2837
Score: 117.50 Matches: 143
Percent Similarity: 33.38% Conservative: 111
Best Local Similarity: 18.79% Mismatches: 252
Query Match: 2.49% Indels: 256
DB: 11 Gaps: 35
US-09-914-168-2 (1-919) x BC032614 (1-2837)
QY 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleLeuAsnHis 41
    |||:|||||:|||||:
Db 1099 CCCATGATAGCTCGGAGAGCTC-----CCATGCCCACTGACAGGCTG 1143
QY 42 ValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeu 61
    |||:|||||:|||||:
Db 1144 GGCCCCCTCAGAGGAGAAAGATGACAGCGAGGCGAGCTCCATCCCTTCACAGAGTGC 1203
QY 62 ThrProGlnGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSer 81
    |||:|||||:|||||:
Db 1204 CCCGCCCGCCACCTGCTCGCAAGAAAGAAACCTTCCC-----ACTGCCCTCCCAAGAGCGC 1257
QY 82 GlnAlaLeuAspValValAsnPhaAspAspGlnSerPro---IleSerArgIleGlyGlu 100
    |||:|||||:|||||:
Db 1258 CGCGTTTCGAGAGGAGTCTCTTAGAAGACCAAAAGTCCGGGATGGCGGACAGAGGGGAC 1317
QY 101 Gln-----SerProProLeuGlyLeuAspMetSerValIleGlnGluThrThrProLeu 118
    |||:|||||:|||||:
Db 1318 CAGCTCAGCCGTGCTCCCAAGGG-----ACCTCAGAGGGCCCTGAGAGACACACC--- 1368
QY 119 SerLeuGlnGluLeuPheAlaGlnGlnSerThrGlnMetGly-----IleAsn 134
    |||:|||||:|||||:

```

Db 1369 -----CGGAGAGCAGCAGCAAGGCCAGACAGAGGTGAAA 1407  
 Oy 135 ProasnaSPYRIleProGluTYrGlnGluGlnProasnserserGlnValValPro 154  
 Db 1408 GCCAGCGAT-----CCTCACAGCATGCCAGAGCTGCCAGAGACA---GCCAACAACACC 1458  
 Oy 155 ProThrleuGluProGluLysProGluLeuLeuLysArgLeuTYrAlaArgLeuPheasn 174  
 Db 1459 CCAGTCCCGCCCGCCAGAAAAACGATCTCTCGACAACTGGCCTCGACCTC----- 1512  
 Oy 175 AspGlyValAsnLysValProArgLeuLysAlaLysPheTYrGlnserSerGlnserGly 194  
 Db 1513 -----CCAGCTCTCTAGAGAACGCTGAGCTCTGACACAGCAGCGATGGCTTG 1560  
 Oy 195 GlnThr-----SerAlaileGlyser 201  
 Db 1561 GAGACACCCAGCCCGGCTCCACCCAGAGAGGCCAAAGCCCTGCTTCCAGGCTGGACT 1620  
 Oy 202 SerHisGlnLysThrGluProTYrAlaAsnileLysAlaAlaLeuGluAspIleThrGln 221  
 Db 1621 CAGCACCTCTCCCGCCAGCCACATGCCAT-----TCC 1653  
 Oy 222 GluseralmetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAla 241  
 Db 1654 CAGAGCTCTCCAGAGTTCAAGGCTCC----- 1680  
 Oy 242 AlaArgAlaValGlyTYrTYrAspIleAspLeuSerIleleArgnserIleGlyGlu 261  
 Db 1681 -----CTGGCTCCCTCTCGACAGCTGGGG--- 1707  
 Oy 262 ValAspValIleIleHisAspLeuGlyGluProValTYrIleAspTYrArgAlaValGlu 281  
 Db 1708 GTGCTGTCAATGAGCCAGCCAGCAG-----GACTCTACTCCAGCCAGC 1749  
 Oy 282 ValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProleu 301  
 Db 1750 AGCAGCCAGCGAGAGCTGGAGCTTCAGCAGCCAGCCAGCTGAGACAGACAGCCCTCCACTG 1809  
 Oy 302 LeuIleGlyAspValPheHisIleGlyLysTYrGluThrLysLysAsnLeuIleGluAsn 321  
 Db 1810 ATCTGGGCAAGGCTCGCGAGCTGAGCTTGGCACTTTCAGCAGCATGTC----- 1863  
 Oy 322 AlaSerAlaGluHisGlyTYrPhe---AspGlyArgTYrPheAspArgSerValAspVal 340  
 Db 1864 -----CACGCTTCTCTCCAAACAGCAGCAGCTGTACAGAAAGTG----- 1905  
 Oy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTYrAspThrGlyThrGlnTYr--- 359  
 Db 1906 -----CTGAGCTGGCCGACGACAGCGCTCTGACTTTGGC 1941  
 Oy 360 ---ArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThr 378  
 Db 1942 AGCGTGTGCGAGACTACAGCTGTACAGCTTGAGATGATGAGCGCGCAG---ACCTCC 1998  
 Oy 379 AspProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMet 398  
 Db 1999 AGCAGCGAGATGCTGACAGAGATTGCGACATGATGACCCAGCTC----- 2043  
 Oy 399 GlyGluAlaTYrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418  
 Db 2044 ---AAGACTCTACTGCTCAG----- 2061  
 Oy 419 TYrPheAsnMetValAsnThrGluIleValPheProGluArgGlnIleGlnAsnAsp 438  
 Db 2062 -----AGCAGCGAGCTC----- 2073  
 Oy 439 GlnValSerPheGluGlnSerSerSerArgTYrGluProAlaGlnValAspGluSer 458  
 Db 2074 -----AAGCCCTGTGTGAGCCCGCTCCACTCCGACGAG 2109  
 Oy 459 ThrLeuGluProValIleGluThrVal-----Glu 468  
 Db 2110 GAGCTCCAGACAAATTGTAGAGTCTGCTTGTACAATGTGTCTGAAGCCCTCAGAGAA 2169

Oy 469 LeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIle 488  
 Db 2170 GCCATCAACTCATGCTCTGCATCATGATCCACAGCAAGATGTTGCTGCTGACAGCTCAAG 2229  
 Oy 489 GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTYrAspMetProAspAsp 508  
 Db 2230 GAGAACCATGATTAGTACTGCTGGCCACCCACCATCATGACTCA----- 2271  
 Oy 509 ArgValIleuAlaIleAsnHisAspAspGlyValAsnArgSerIle----- 523  
 Db 2272 -----GGTGGACCACCGCTGCTCGGAGTGGCCATG 2304  
 Oy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543  
 Db 2305 ATGAGAAAGATCCCTGCACAGATTCACAGCATGACAAAGGCTTACTACCTGAGAAAGAG 2364  
 Oy 544 GluAsnGluValIleAsp-----LeuProGluArgTYrAlaLeuAlaAsnArg 559  
 Db 2365 ATCTCATCTGCTTCACAGCTGCAAACTCATCTGAGCTCCATGCGCCGCAACCA 2424  
 Oy 560 LysThrProAlaAspValTYrGlnserLysValProleuTYrValPheValAlaSer 579  
 Db 2425 GGAAGCCCTATGGCGGAGTACTTCTGCTGCTCATGATGATGCTGGCCGCGCAGC 2484  
 Oy 580 Asp-----LysProArgAspGlyGlnIle 587  
 Db 2485 AACCTCAGCGAGATCTTCTCAATGTGAGTACATGATGAGCTCATGAGCCCGCCCTG 2544  
 Oy 588 GlyLeuGlyTYrPglySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeu 607  
 Db 2545 CAGCTGGGAGGAGGTTCCTACTATCTGACACACCTACGGGGCCCTGAGCAGC----- 2598  
 Oy 608 IleAsnArgAspGlyTYrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGly 627  
 Db 2598 ----- 2598  
 Oy 628 ValLysLeuTYrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThr 647  
 Db 2599 ATCAAGACTTAC-----GACAAATCATCAGGTG--- 2625  
 Oy 648 LeuGlyTYrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg 667  
 Db 2626 -----ACCCGG 2631  
 Oy 668 ThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTYrAsnArgThrTYr 687  
 Db 2632 CAGCTGACTGTGAGGTGCAGAGTCAATCCACGCC-----TGGGAGGCGCGGCT 2682  
 Oy 688 SerLeuArgTYrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTYrGlnAsp 707  
 Db 2683 ACTCTCAAC-----AAGGCCCGGCTCTCCGCTCTCCGTACAGAGAC 2724  
 Oy 708 ---LeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726  
 Db 2725 TTCATCTCGTGTCTTACTG-----GAGCCGAGCAGCAGCGCGGAGC-GCTGGCGTC 2777  
 Oy 727 Ala 727  
 Db 2778 GCG 2780  
 RESULT 13  
 AV557254/c 607 bp mRNA linear EST 06-SEP-2000  
 LOCUS AV557254 Arabidopsis thaliana green siliques Columbia Arabidopsis  
 DEFINITION thaliana CDNA clone S0064409f 3', mRNA sequence.  
 ACCESSION AV557254  
 VERSION AV557254.1 GI:8728669  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;





ACCESSION	melanogaster cDNA clone LD39445 5prime similar to M31617; trx					
VERSION	F80n0003662 P1D:g156818 SWISS-PROT:P20659, mRNA sequence.					
KEYWORDS	A1519730					
SOURCE	A1519730.1 GI:4425584					
ORGANISM	EST.					
TITLE	fruit fly.					
AUTHORS	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pleiygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridoidae; Drosophilidae; Drosophila. 1 (bases 1 to 544) Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M. BDGP/HHMT Drosophila EST Project Unpublished (2001)					
JOURNAL	Contact: Stapleton, M. BDGP					
COMMENT	Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6796 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu Plate: 394 row: D column: 9 High quality sequence stop: 543. Location/Qualifiers 1..544 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_ldb="LD39445" /cgene_ldb="LD Drosophila melanogaster embryo POT2" /sex="male and female" /dev_stage="0 to 24 hours mixed stage embryonic" /lab_host="XLI Blue" /note="Organ: embryo; Vector: pOT2; Site:1; EcoRI; Site_2 XhoI; Sized fractionated cdnas were directly ligated into pOT2."					
BASE COUNT	155 a 161 c 118 g 110 t					
ORIGIN						
Alignment Scores:						
Pred. NO.:	0.0155 Length: 544					
Score:	112.00 Matches: 39					
Percent Similarity:	45.16% Conservative: 31					
Best local Similarity:	25.16% Mismatches: 63					
Query Match:	2.37% Indels: 22					
DB:	Gaps: 9 7					
US-09-114-168-2 (1-919) x A1519730 (1-544)						
Oy	13 MetProValAlaLeuAlaTYrLeuProMetThrSer----- 26					
Db	84 ATGCCCATTTATTAACCTTGCTGCAGCCACCAGTAGTCATCCCAATTGTGTACGACCT 143					
Oy	27 GlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisasp 46					
Db	144 CAACTTTTTGCAGCAGCAG---CAGCTGGCTAACCCCGTAGCAGCATTTCTCCACAAGCAGC 200					
Oy	47 ThrAlaIleasnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluInile 66					
Db	201 AGCACT---AGCAGTAGCACACTGCTCACTGCCACCAATGTGGTTAAATCTATGCACAACA 257					
Oy	67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86					
Db	258 CAAGCACCATCAACCAAGCAGCAGCTCCACAACCAAGCCTTACAAATCGGGCTGCTACCGATG 317					
Oy	87 ValAsnPhesaspASpGlnSerProIleSerArg-----IleGlyGluGlnSer 102					
Db	318 CAGCAGCGCCCAAGGACCTGCTCGTTGTCACAAGCAGTCCCGTAGTTTCGTACCCCACT 377					
Oy	103 ProProuLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGlu 122					
Db	378 CTCGCGAANACCTGTTGAACAGCGCTTATTCATCAATAATGACTACAGCCGCAAGTGTGTCAAG 437					
Oy	123 LeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrlIleProGluTyrr 142					

[illegible]

```

Db 1005 ACTTCCTGTCACCGAACAAGCTGCAGAAATTAAAGATTCAGATACAGACGAGNAATT 1064
Oy 464 egIuThr-----ValGluLeuThraspgIyle-----LeuMetAspTleserProII 480
Db 1065 GGAAGGCTTGTTGAGGTAGCAAGGTGAATGGGTTGCTCTTAATTATTAAGTCCAGCA 1124
Oy 480 egIuPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaAr 500
Db 1125 AGAAGCTGGTCTTGCAGAGCTTACAGGAATCATGTGCACTATACACAGCTGGCAACAG 1184
Oy 500 gHISLeuTYrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspIyValAs 520
Db 1185 TCACATGCTTGA-----CTCAGCATG-----CAGATGTACAGGTGGA 1223
Oy 520 naGserIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPr 540
Db 1224 CAATCAGCTCATTTGTAACCAAGCCCTTCATGCTCTATGTGACT-----CC 1271
Oy 540 oAspGluSerGluAsnGluValIleAsp----- 549
Db 1272 CCTGAGACATGAGAAATGAGGTGATCGAGCGCCAGCTGTGCAAGTCAACGCAATGAA 1331
Oy 550 -LeuProGluArgThrAlaLeuAlaAsn-----Argly 560
Db 1332 GTTCCCCAGTAAAGAGTGCACCTCAACATCTACAAACATCTGATGATCAGCTCAGAG 1391
Oy 560 sThrProAlaAspValTYrGlnSerLysValProLeuTYrValPheValAlaSerAs 580
Db 1392 ATTCAAGTGCACAAATTGAGAGAAACTGCTCTCAAGCTCTTAAGTTTC----- 1440
Oy 580 pLysProArgAspGlyGlnIleGlyLeuGlyTYrGlySerAspThrGlyThrArgLeuVa 600
Db 1441 -----TTGGCTAGACATCAAGCAATCAGAGT 1469
Oy 600 lThrLysPheGluHisAsnLeuIleAsnArgAspGlyTYrGlnAlaGlylaGluLeuAr 620
Db 1470 GGAATAATATGATGAATAAACCCTCATGAAGAAGACAGCTGAGCAAGTGCACCAATTCG 1529
Oy 620 g-----LeuSerGluAspLysGlyValLysLeuTYrAlaThrLysPr 635
Db 1530 ATTACTACTTTGAAAATCTCAAAATCAAGCATCTCTCAGATCAAGCTAGTGTTCACCTC 1589
Oy 635 oLeuSerHisProLeuAsn---AspGlnLeuArgAlaThrLeuGlyTYrGlnGlnGluVa 654
Db 1590 CAACAAGCTCCCATGTGATCTTAAGGCCCTAAAGACGACCTGGGTTCTTGATACG 1649
Oy 654 lPheGlyHisSerThrasnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSe 674
Db 1650 GTTTGAAGACGCTGTGATTAATCTAGATTCATTCACTCGGGTACATCCCTATGAGACCA 1709
Oy 674 rArgSerIleIleGlnAsnGlyTYrPAsnArgThrTYrSerLeuArgTYrArgLeuAs 694
Db 1710 GAGGTTCAATCATCAAT----- 1725
Oy 694 pLysLeuLysThrGlnAlaProProGluThrTYrGlnAspLeuProValAspPheValAs 714
Db 1726 -----GATATCCTC-- 1734
Oy 714 nGlyLysProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAs 734
Db 1735 ----AAACATTTCCAGAGAACTCTCAGCCAGCGAGCTGCAATCCTGGATCAGATGGA 1790
Oy 734 pAsnLeuValAsnProMetArgGlyTYrArgGlnArgTYrSerLeuGluValGlySerSe 754
Db 1791 TTTTCTTGGCAATCCTATG----- 1809
Oy 754 rGlyLeuValSerAspAlaAsnMetalAlaAlaArgAlaGlyIleSerGlyValTYrSe 774
Db 1810 -GGGCTTTGATGATGATTTCT-----GAAGGGGTACTGAGACTGATGAAA 1853
Oy 774 rPheGlyAspAsnAlaTYrGly-----SerAsnArgAl 785
Db 1854 ATATGGA---AATGTGGGGGCTCATCAGAAATGTTACACAGGAGTATCAAACTCTGC 1910

```

```

Oy 785 aHISGlnMetThrGlyGlyIleGlnAlaGlyTYrIleTrpSerAspAsnPheAsnHisVa 805
Db 1911 TGGCAATTTGCTGGAAACATTATCAGATGGCTTAGGAAACAGCATGACATCGGCATCA 1970
Oy 805 lProTYrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgly 821
Db 1971 GTCAGACCG---GATACATCATGAGTGAAGGTGGAACAAAGAGCGCG 2016

```

```

RESULT 17
BF491509 678 bp mRNA linear EST 19-Apr-2001
LOCUS AT8265.5prlme AT Drosophila melanogaster adult testes potB7
DEFINITION Drosophila melanogaster cDNA clone AT8265 5 similar to trx:
FBan008651 'transcription factor' located on: 3R 88B5-88B6::
04/09/2001, mRNA sequence.
ACCESSION BF491509
VERSION BF491509.2 GI:13695035
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 678)
AUTHORS Stapleton,M., Broksstein,P., Hong,L., Agbayani,A., Baxter,E., Berman
,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V., Farfan
,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P.,
Liao,G., Miranda,A., Mista,S., Mungall,C.J., Nunoo,J., Pacled,J.,
Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E.,
Celinkier,S. and Rubin,G.M.
BDGP/HIMI AT Drosophila EST project
JOURNAL Unpublished (2000)
COMMENT On Dec 6, 2000 this sequence version replaced gi:11574810.
CONTACT: Stapleton, M.

```

```

FEATURES
source
1..678
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="AT8265"
/clone_lib="AT Drosophila melanogaster adult testes potB7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="plates AT.10-AT.120: DHS-alpha. plates
AT.121-AT.319: DHS-alpha Tona"
/note="Organ: ADULT testes; Vector: potB7; Site.1: ECoRI;
Site.2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into potB7. Plasmid cDNA library."

```

```

BASE COUNT 201 a 178 c 166 g 133 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 0.0346 Length: 678
Score: 110.50 Matches: 48
Percent Similarity: 41.82% Conservative: 44
Best Local Similarity: 21.82% Mismatches: 83
Query Match: 2.34% Indels: 45
DB: 12 Gaps: 11

```

```

US-09-914-168-2 (1-919) x BF491509 (1-678)

```



```

QY 203 ---HisGlnLysThrGlu-----ProTyrAlaSerIleLysAlaIleGlu 217
      ||||| |||
Db 824 ATACACACATGAGAGAGAACACACACAGTACAGAGACTCCAGGCAACCCCTGCCAA 883
QY 218 AspIleThr-----GlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArg 235
      ||||| |||||
Db 884 GAGCTAGCTGATTACAGACAGATTACCCAGACACTGATAGTGAACGAAGGCTTGGG 943
QY 236 GlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIle 255
      |||||
Db 944 GAGAGAGAGGTTATT-----CTGATG 964
QY 256 ArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIle 275
      ||||| |||||
Db 965 GAGTCTTTATGTCAGACAGACCATAGTTCGACACTTATAGTCGACAG-----ATT 1015
QY 276 AspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrVal 295
      |||||
Db 1016 GAATAC-----TTCCGCTCTCTT 1033
QY 296 AlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLys 315
      |||||
Db 1034 CTAGATGAG-----CATCACATTTCTTATGTCATAGAT 1066
QY 316 LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPleuAsp 335
      |||||
Db 1067 GAAGATGTA-----AAAGTGGCGCTATATGGAATTAGACAGACTTATCATGTCGAC 1117
QY 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355
      |||||
Db 1118 -----CTCCGCTGAGAAATGCC----- 1132
QY 356 GlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGln 375
      |||||
Db 1133 -----CGTTTGAACG----- 1144
QY 376 LeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeuLeu--- 394
      |||||
Db 1145 -----GAGCAGCTTCTTGGT 1159
QY 395 -----ThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaVal 407
      |||||
Db 1160 GTCCACACACATTTAAGCAATATCTTGAATAATGGCAGACACACAGCAATAGGAAGCTCAA 1219
QY 408 ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIle 427
      |||||
Db 1220 GAATGATAGCG-----GCACCTCAAGAACGACGATCCCATATG-----GAGCGAATT 1267
QY 428 ValPheProGluArgGlu----- 433
      |||||
Db 1268 ATTGAATCTGAGCAGAAAGAAAGAACAGCGCTTGGCAGCCACGTTAGAGATCAAGGCC 1327
QY 434 GlnIleGlnAsnAspGlnValSerPheGlu-----GlnSerSerSerArg 449
      |||||
Db 1328 ACAGTGGCCAGTCGACAGATAGATGAATCGCCTGAGAGCTCGAGATGAAG 1387
QY 450 ThrGluProAlaGlnVal-----AspGluSerThrLeuGluPro 462
      |||||
Db 1388 CAGAAAGTGGCAGAGCTGATTTATTCATTACTCTGGAGACAAATCTGATATTACGAGAC 1447
QY 463 ValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPhe 482
      |||||
Db 1448 CTCCTGAGAGTGTGAGGCG-----GACAAAGAAAGAGAGAGACT 1489
QY 483 SerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeu 502
      |||||
Db 1490 TTGGCTAGTAGCTTGCAGGAGAT----- 1513
QY 503 TyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSer 522
      |||||
Db 1514 -----CTGGCTCATACCCGGAATGATGCG----- 1537

```

```

QY 523 IleLeuGlyArgIleSerAspAlaValSerAlaValAla-----ArgAlaIleLeu 539
      |||||
Db 1538 -----AATCGATTACAGATGCGCTTGGTATAGTAGAGATCAATACCGAGC-----TTC 1588
QY 540 ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
      |||||
Db 1589 CAAGAAGAGAGCTAAGAAACAAATTTGAAGATTG-----AAATAGAGCTTAGAA 1636
QY 560 LysThrProAlaAspValTyrGlnSerLys 570
      |||||
Db 1637 AAATTAAGATCAGACCTGGATGAAAAA 1669

RESULT 19
AY104509 1501 bp mRNA linear HTC 25-MAY-2002
LOCUS AY104509
DEFINITION Zea mays PCO104742 mRNA sequence.
ACCESSION AY104509
VERSION AY104509.1 GI:21207587
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1501)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
Arthur,L.W., Hanfey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/Dupont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 1501)
Coe,E.C.
TITLE Direct Substitution
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
FEATURES
source
1..1501
/organism="Zea mays"
/db_xref="MazEDB:636055"
/db_xref="taxon:4577"
/clone="PCO104742"
/clone_lib="Maize Mapping Project/Dupont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT 308 a 489 c 418 g 286 t
ORIGIN
Alignment Scores:
Pred. No.: 0.215 Length: 1501
Score: 109.00 Matches: 98
Percent Similarity: 31.26% Conservative: 53
Best Local Similarity: 20.29% Mismatches: 168
Query Match: 2.31% Indels: 164
DB: 11 Gaps: 23

US-09-914-168-2 (1-919) x AY104509 (1-1501)
QY 439 GlnValSerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSer 458
      |||||
Db 149 GAATCTATCTCGCTCGCGGCTTCACTCACTCAATTTCCGACACAGTTCGACAG 208
QY 459 ThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSer 478
      |||||
Db 209 AATACCGACCA-----AGGACGACGACGAGCGTCGATGGCAGCGTCTCC 253
QY 479 ProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLys 498
      |||||
Db 254 -----TCCTGCACCGCGCTGAGTAGCGGC-----ACCGCGCGCGG 289

```





AUTHORS Walbot, V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 614073 row: B column: 08.  
Location/Qualifiers

FEATURES  
source 1..588  
/organism="Zea mays"  
/cultivar="W23"  
/db\_xref="taxon:4577"  
/clone\_lib="614 - root cDNA library from Walbot Lab"  
/tissue\_type="root"  
/dev\_stage="3-4 days old"  
/lab\_host="XLOLR"  
/note="Organ: root; Vector: pBluescriptII SK+; Site: 1; EcoRI; Site: 2; XhoI; 3-4 days old root tissue from Walbot Lab (Lm)"  
BASE COUNT 182 a 156 c 112 g 138 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.0939 Length: 588  
Score: 106.00 Matches: 34  
Percent Similarity: 41.60% Conservative: 18  
Best Local Similarity: 27.20% Mismatches: 41  
Query Match: 2.24% Indels: 32  
DB: 10 Gaps: 5

US-09-914-168-2 (1-919) x AM053211 (1-588)

OY 788 MetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPhaSnHsValProTyr 807  
:::|||||  
DB 543 TTGACAGCGCGATCCATGTTGTCG-----GACATGCGCCTTAT 505

OY 808 ArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrPheAlaHisAspSerLeu 827  
:::|::|  
DB 504 CAA---GCTTTTGCCATGAGTGGCTGAGTCCGAGGCTATGCT----- 460

OY 828 SerProIleSerAspGlyGlyTyrLeuThrGlyGlnValLeuAlaValGlyThrAla 847  
:::|::|  
DB 459 -----GAGGGCGCTGTTGCTTCAGAAAGACTATGCTAAATTGCTAACTGC 415

OY 848 GluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsn 867  
|||||  
DB 414 GAATATACGATTCCTCTTGGAAGCATCTGAGAGGTTCTATTTCATGTGACTGTGGAGT 355

OY 868 AlaTyrAspGlyPheThrAsnAspThrLys----- 878  
|||  
DB 354 -----GACTTAGGGTCTGCTGCTCATGTCCTGTAATCCAGCTCGCGCAGGGGAAA 301

OY 879 -----IleGlyAlaGlyValGlyValArgTyrPheLaserProValGlyGlnVal 894  
|||||  
DB 300 CCAGGATTTGCTATGATTTGGATACGAGACTTCAACACTGACCTAGGTCAAGATC 241

OY 895 ArgValAspValAla 899  
|||||  
DB 240 CGTGTGACTACGCT 226

RESULT 22  
LOCUS BE788630 758 bp mRNA linear EST 20-OCT-2000  
DEFINITION 601475986F1 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:3878925 5',  
mRNA sequence.  
ACCESSION BE788630  
VERSION BE788630.1 GI:10209828  
KEYWORDS EST.

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 758)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: DCTD/DRP/Gazdar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
Plate: LLM9643 row: k column: 22  
High quality sequence stop: 339.  
Location/Qualifiers

FEATURES  
source 1..758  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:3878925"  
/clone\_lib="NIH\_MGC\_68"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by life technologies."  
BASE COUNT 212 a 214 c 235 g 97 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.194 Length: 758  
Score: 105.00 Matches: 53  
Percent Similarity: 42.64% Conservative: 31  
Best Local Similarity: 26.90% Mismatches: 85  
Query Match: 2.22% Indels: 29  
DB: 12 Gaps: 8

US-09-914-168-2 (1-919) x BE788630 (1-758)

OY 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540  
|||  
DB 59 CGGAAGTGTCTC-----ATCAGTGACAGCGCTTGACCTTGTCCGGAAGATCTTGCCA 112

OY 541 AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560  
|||  
DB 113 GATGGAGG-----CTGCAGGTGCTGGAAAGACAGAACCTTGACAAAGAGAG 160

OY 561 ThrProAlaAsp-----ValTyrGlnSerLysLysValPro 572  
|||  
DB 161 CTGATAGCGGAGCTGCGAGAGCTGGAAGGCTTATGTTGCTGCTGCCACCAAGGTGACC 220

OY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGly----- 590  
|||  
DB 221 GCTGATGTCATCAACGAGCTGGAACCTCCAGGTGTGGCAGCGGCTGGCAGAGT 280

OY 591 -----TrrGlySerAspThrGlyThrArgLeuValThrLysPheGluHis 605  
|||  
DB 281 TGGACATGTGATGTCGAGCGGCAACAGGA---AGGCACCTCTGTGTTATGAACAC 337

OY 606 AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLys 625  
|||  
DB 338 CCCCAACAGGAGGAGCAAGACCTCAAGATGCGCCGCAACCAACATCAACATAGACT 397

OY 626 LysGly-----ValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
:::|::|  
DB 398 CGGGGCAATGAATCACTAGTCGAGCCATCGGCGCACAGGCGAGG-ATGTAATCCCGGAGC 456

OY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661



```

Db 457 GGGCGACAGAGGCGACTGTCGATAGACAGGACCGGACGACAAA---AATGGG 513
      :::: ||| ::::||| |||
Qy 662 PheAspLeuSerThrArgThrLeuLHisGluIleSerArgSerIleIleGlnsngly 681
      ||||| ||||| |||
Db 514 AGAACCCGAGCAGCAACCCGGATV---GCCACTCGGGGGGCAACGCAACGGG 570
      :::: ||| ::::|||
Qy 682 GlyTPAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuIsth 698
      :::: ||| ::::|||
Db 571 AAGGAGTCTAGACAACTCGGTGACACAGGACGACACCCCTCGGAGC 621
      :::: ||| ::::|||

RESULT 23
BH770729/c 1557 bp DNA linear GSS 01-MAY-2002
LOCUS LMWtag478 MG1363 Random Sequence Tag Library Lactococcus lactis
DEFINITION subsp. cremoris genomic, DNA sequence.
ACCESSION BH770729
VERSION BH770729.1 GI:20373686
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
          Lactococcus.
REFERENCE 1 (bases 1 to 1557)
AUTHORS Bojotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Aliments, (2002) In press
COMMENT Contact: Sorokin A
          Genetique Microbienne
          INRA
          CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
          Tel: 33 1 34 65 25 16
          Fax: 33 1 34 65 25 21
          Email: sorokine@jouy.inra.fr
          best homologue in strain IL1403 is ywfg (76%)
          Class: shotgun
          High quality sequence start: 30
          High quality sequence stop: 1527.
FEATURES
    source
        1..1557
            /organism="Lactococcus lactis subsp. cremoris"
            /strain="MG1363"
            /db_xref="taxon:1359"
            /clone_lib="MG1363 Random Sequence Tag Library"
            /note="Vector: pSGM02; Site 1: SmaI; Library of
            chromosomal fragments of L.lactis strain MG1363 was
            prepared by partial AluI digestion or by sonication."
BASE COUNT 403 a 318 c 253 g 582 t
ORIGIN
Alignment Scores:
Pred. No.: 0.696 Length: 1557
Score: 105.00 Matches: 135
Percent Similarity: 29.24% Conservative: 70
Best Local Similarity: 19.26% Mismatches: 200
Query Match: 2.22% Indels: 296
                Gaps: 34
US-09-914-168-2 (1-919) x BH770729 (1-1557)
Qy 181 ProArgLeuLysAlaLysPheTyrGlnSerGlnSerGlyGluThrSerAlaIleGly 200
      ||||| ||| ||| |||
Db 1544 CCCCCGTTT-----CTCCGTCAGTAACAACANGTCGAGACATTAACACAGCTTAC 1494
      :::: ||| ::::|||
Qy 201 SerSerHisGlnLysThrLysProTyrTrpAlaAsnIleLysAlaIleLysLysPheThr 220
      :::: ||| |||
Db 1493 ACGACAGACCAAAAATATTCCA-----GGATATACT 1461
      :::: ||| |||
Qy 221 GlnGluSerAlaMetAspLysnnglySerIleProArgLeuArgGlnThrAlaLeuVal 240
      :::: ||| |||
Db 1460 TATCATATCATCA-----AATGGA----- 1443
      :::: ||| |||
Qy 241 AlaAlaArgAlaValGlyTyrTyrAspLysLeuSerIle-----IleArg 256
      :::: ||| |||

```

```

Db 1442 GCAGCACAAGCGAAGAACTATACGTTAATGATCAAACTGTAATATGTTATACAGA 1383
      ||||| :::: |||
Qy 257 AsnSerIleGlyGluValAlaAspValIleIleHisAspLeuGlyGluProValTyrIleAsp 276
      ||| :::: |||
Db 1382 AATCAAA---GGAAAGCATGATGTA-----ACCTACATTGAT 1350
      :::: |||
Qy 277 TyrArgAlaValAlaGluValArgGlyLys----- 285
      :::: |||
Db 1349 CAGACAAACGGCGAAGTTTGACGAAGAAAGATTGTCTCGTGAAGTATGACCCCTGG 1290
      :::: |||
Qy 286 -----GlyAlaAsp-----AspLysAlaPheThrThrValAla 296
      ||||| |||
Db 1289 AATTATACAAACAGTCGATATTAATCTTACACGATAAAGATTAATGATTAAGTTCT 1230
      :::: |||
Qy 297 AspGluValProLeuLeuIleGlyAspValPhe-----HisIleGlyLys 311
      ||| :::: |||
Db 1229 GATGATTATCCAGTCAA---GGACGAGTATTTACGGATACCTCCGACGATTATGAGTT 1173
      :::: |||
Qy 312 TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluIleGlyTyrPheAspGly 331
      ||| :::: |||
Db 1172 TATTTAAGCAAAATACAGTTGTAAGTTGGAACAAACAA----- 1131
      :::: |||
Qy 332 ArgTyrLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu 351
      :::: |||
Db 1130 -----GTGAATGAACCATTCATTAT 1110
      :::: |||
Qy 352 IleTyrAspThrGlyThrGlnTyrArgPheAsp-----GluValAlaPhePheThr 368
      :::: |||
Db 1109 GTCTATGAAGATGAAGACATGAAGCTCGACGAGACTATAAAGCAGACCGCTGAATTTGACA 1050
      :::: |||
Qy 369 IleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgLys 388
      ||||| :::: |||
Db 1049 AGAACCGTAAACAAGATCAAGTAACGGGTGAAAA----- 1014
      :::: |||
Qy 389 LeuLeuGluGlnLeuLeuThrValAsnMetGlyAlaIleTyrAsnLeuGlnAlaValArg 408
      ||| :::: |||
Db 1013 -----ACATACGAGAAATGGCAAGCGGAAAT----- 987
      :::: |||
Qy 409 AlaLeuSerAsnAspLeuLeuLeuThrArgTyrPheAsnMetValAsnThrGluIleVal 428
      ||| :::: |||
Db 986 -----GGTAATAGTTT-----CTTAAGTAGTT 963
      :::: |||
Qy 429 PheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSer 448
      ||| :::: |||
Db 962 TCACCAAGTATGAAGAGTGAACACCGACGCA----- 930
      :::: |||
Qy 449 ArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGlu 468
      ||||| :::: |||
Db 929 -----GCTCAATGATGAAATACG----- 909
      :::: |||
Qy 469 LeuThrAspGlyIleLeuMetAspLysSerProIleGluPheSerAlaSerAsnLeuIle 488
      ||| :::: |||
Db 908 -----GGTGTACACGACAGATACGACGATTTAGG----- 879
      :::: |||
Qy 489 GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp 508
      :::: |||
Db 878 -----AAA 876
      :::: |||
Qy 509 ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528
      :::: |||
Db 875 AAGATATCTATTAACGATATCAAGGACGATGATGATCTATATTT----- 828
      :::: |||
Qy 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluIle 548
      :::: |||
Db 827 -----GATGAACGACAGTCAAGTTCTG 804
      :::: |||
Qy 549 -----AspLeuProGluArgThrAlaLeuAlaAsnArgIsthThrProAlaAspVal 565
      ||||| |||
Db 803 ACTAAAAAGATTTTCTGCTGGTGAACGATGATCCGCTCAATATATGCAACAGGTGAGAT 744
      :::: |||
Qy 566 TyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGly 585
      :::: |||

```

```

Db 743 ATTAAATCTTACACAGATTAAGCGTTACGAATTAGTTCAAGTGATTAATCCAGATGGCGG 684
Oy 586 GllleGlyLeuclYtrpGlySerAspThnGlyThrArgLeuValThrlYspheGluHis 605
Db 683 ACATC-----TTTACTGATGAAGCACAACATTATACCGTGAAATTGAAACAT 636
Oy 606 AsnLeuLeuAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLys 625
Db 635 GGATTAACAAT-----GAACTGATTA 612
Oy 626 LysGlyValLysLeu-----TyrAlaThrLysProLeuSerHisProLeuAsnAsp 642
Db 611 AAGCTGTCAATCAAGTATTCATTATGTTTATTAAGATTAAGTCAAGATAAGCAGCAGAG 552
Oy 643 GlnLeuAlaGlnThrLeuGlyTyrGlnGlnGlnValPheGlyLysSerThrAsnGlyPhe 662
Db 551 GACCATACGCTACTCTTGATTTAGTCACAATTA----- 516
Oy 663 AspLeuSerThrArgThrLeuGlnHisGluLysSerArgSerIleIleGlnAsnGlyGly 682
Db 515 -----AGTACTGACAACAGTAACAAAAG----- 492
Oy 683 TTPAsnArgThrSerLeuArgTyrArgLeuAspLysLeuYsthrGlnAlaProPro 702
Db 491 -----AAACTTATAGC----- 480
Oy 703 GluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnAlaLeu 722
Db 479 -----GACTGC----- 474
Oy 723 LeuAlaGlyValAlaValHisLysThrValAlaAspAsn-----Leu 736
Db 473 -----GCGCGTCAACAATGCTGACAGCGTTTCCGAAGTC 441
Oy 737 ValAsnPro---MetArgGlyTyrArgGlnArgTyrSerLeuGlnValAlaGlySerSerGly 755
Db 440 GTTCCCGCACTGCTCAAGAGTTTATACAGCTGACCAACTTAAGTCTGTGAATGACAGCA 381
Oy 756 LeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyrSerPhe 775
Db 380 ATCATTCTGCACACA-----GCAGATATTCTCAGTAACGCGTTTCTTAT 339
Oy 776 GlyAspAsnAlaIleTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGly 795
Db 338 ACAAGAAAT-----CAAGGAACGATTGATGTCGACT 309
Oy 796 TyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGly 815
Db 308 TATATT-----CATGAACCACTGGGAAGCTTATCTGCAAAAGATTGTCAGCGCGA 255
Oy 816 AspGlnSerIleArgGlyTyrAlaHis---AspSerLeuSerProIleSerAspLysGly 834
Db 254 ACTGAGATTTCTTCAACACTATACACGCTTGATACGATTAAATCTTACACAGATAAGGCT 195
Oy 835 Tyr 835
Db 194 TAC 192

RESULT 24
CNS071R5/ 1021 bp DNA linear GSS 07-JUL-2001
LOCUS BAOAB004F03 of library BAOAB from strain CLIB 210 of
DEFINITION Kluveromyces lactis, genomic survey sequence.
ACCESSION AL425271
VERSION AL425271.1 GI:12208465
KEYWORDS GSS.
SOURCE Kluveromyces lactis.
ORGANISM Kluveromyces lactis
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Kluveromyces.
REFERENCE Soulel,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
COMMENT
FEATURES
source
misc_feature
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-09-914-168-2 (1-919) x CNS071R5 (1-1021)
Oy 424 AsnThrGluLeuValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGlu 443
Db 744 AATTGGCATCTGACCGCATTAAGAAAAAATAGGCCAAGAACAAAGAAAGGCGCACAC 685
Oy 444 GlnSerSerSerArgTrpGlnProAlaGlnValAspGluSerThrLeuGluProVal 463
Db 684 AGTAACAATAAATCAAAAGTCTGCTGTCAGCAAAAGATGTCAAGGCTACAGAGCGAGTC 625
Oy 464 IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSer 483
Db 624 CAAGAGCTGGTCTCGAATAATGATGCG----- 598
Oy 484 AlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyr 503
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1..A set of
yeast species for molecular evolution studies
FEMS Lett. 487 (1), 3-12 (2000)
2 (bases 1 to 1021)
Toffano-Nioche,C., Artiguenave,F.,
Duchateau-Nguyen,G., Lemaire,M., Marmisè,R., Montrocher,R.,
Robert,C., Termler,M., Wincker,P. and Wesolowski-Louvel,M.
Genomic exploration of the hemiascomycetous yeasts: 11.
Kluveromyces lactis
FEMS Lett. 487 (1), 66-70 (2000)
20584721
11152876
11152866
3 (bases 1 to 1021)
Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :
secref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluveromyces thermotolerans, Kluveromyces
lactis var. lactis, Kluveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
Location/Qualifiers
1..1021
/organism="Kluveromyces lactis"
/strain="CLIB 210"
/variety="lactis"
/db_xref="taxon:28985"
/clone.lib="BA0AB004F03"
/clone.lib="BA0AB"
complement(1..>618)
note="similar to Saccharomyces cerevisiae ORF YOR216c [
RUD3 : suppressor of us01-1 transport defect ]"
/evidence="not_experimental"
BASE COUNT 255 a 209 c 184 g 372 t 1 others
ORIGIN

```

Db 597 -----GCTAACGCTGTCCG----- 583  
QY 504 AsmpetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523  
Db 582 ---GTACCGGATGAT-----TCTTAGAAAACTCTGATGGTGGTGCACCACTCAACT 535  
QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspIle 543  
Db 534 GTG-----GAAGCGGTGTACTGTGAAGAACGATGACAGCATCAAACT 487  
QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg----- 559  
Db 486 GACCCAGAGAACACCAATCAACAGCAACCAAACTTAATGACAGGACAGCAACCG 427  
QY 560 -----LysThrProAlaAspValTyrGln 567  
Db 426 CACAGTCATTCTGAGGAGACAGATCAACGCTGTTAAAGAGAAAGATAGTACGAA 367  
QY 568 SerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyIle 587  
Db 366 TTGAAGTATACACCTCTTTGAGTAAATCTCATCTATGAGACGCTGTTCAGTCAATG 307  
QY 588 GlyLeuGlyTyrGlySerAspThrGlyThrArg---LeuValThrLysPheGluHisAsn 606  
Db 306 AAGCAAGCTCAACAGCAAGTGGAAAAACACTCTGTGACAGTTGAATGAGATGAGATCAG 247  
QY 607 LeuIleAsnArgAspGlyTyrGlnIleGlyAlaGluLeuArgLeuSerGluAspLys 626  
Db 246 AATATTAATTTTAAACAGACAGTAAATCTGCTGATTCAGAAATCCAAAGAGCAAGCA 187  
QY 627 GlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspIleuValGla 646  
Db 186 ACATTTCTATT-----TTGAATGATGAAATC---GTC 157  
QY 647 ThrLeuGlyTyrGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666  
Db 156 AATTCATTAAGAAATGTAATTTGAGTATGATGATGATGATGATGATGATGATGATGAT 97  
QY 667 ArgThrLeuGluHisGluIleSerArgSerIleGlnAsnGlyGlyTyrAsnArgThr 686  
Db 96 AAAGATTTGGAATCTTAAGTTGACAGAGCA-----TCAACGAGAGAGTCTAACATTACT 43  
QY 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699  
Db 42 TCTTGTCTTAATCTGAAGTCAAAAGCTTAAATCTCAG 4  
RESULT 25  
AM342205 1364 bp mRNA linear EST 01-JAN-2001  
LOCUS GtHEST91 Gulliardia theta Lambda Zap II cDNA library Gulliardia  
DEFINITION theta cDNA clone Est25\_39\_46\_8, mRNA sequence.  
ACCESSION AM342205  
VERSION AM342205.1 GI:12000580  
KEYWORDS EST.  
SOURCE Gulliardia theta.  
ORGANISM Gulliardia theta.  
REFERENCE 1 (bases 1 to 1364)  
AUTHORS Fruhnholz,M., Duebel,J., Wastl,J., Zauner,S. and Maier,U.-G.  
TITLE EST Database of the cryptomonad alga: Gulliardia theta  
JOURNAL Unpublished (2000)  
COMMENT Contact: Maier, U.-G.  
Department of Cell Biology and Applied Botany  
Philipps-University Marburg  
Karl-von-Frisch-Strasse, D-35043 Marburg, Germany  
Tel: +49 6421 282 2057  
Fax: +49 6421 282 1543  
Email: maiermaier.uni-marburg.de.  
FEATURES  
source 1..1364  
location/Qualifiers  
/organism="Gulliardia theta"  
/strain="CCMP327"  
/db\_xref="taxon:55529"

/clone="Est25\_39\_46\_8"  
/clone\_id="Gulliardia theta Lambda Zap II cDNA library"  
/lab\_host="SOLR"  
/note="Vector: Lambda Zap II; Site\_1: EcoRI"  
BASE COUNT 383 a 286 c 343 g 348 t 4 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.1 Length: 1364  
Score: 102.50 Matches: 71  
Percent Similarity: 36.66% Conservative: 43  
Best Local Similarity: 22.83% Mismatches: 98  
Query Match: 2.17% Indels: 99  
DB: 10 Gaps: 15  
US-09-914-168-2 (1-919) x AM342205 (1-1364)  
QY 261 GluValAspValIleIleHisAspLeuGlyIleProValTyrIleAspTyrArgAlaVal 280  
Db 168 GACGTTGATACATATCATCATCAAGGATTTG-----GTAGACATTCCTGTAGTG 212  
QY 281 GluValArgGlyGluGlyAlaAspAsp---LysAlaPheThrThrValAlaAspGlu--- 298  
Db 213 CTGACGCGGACACAGATGAGAGCAATTTGAAGAGGAGTGGCATGCTGTGGAAC 272  
QY 299 -----ValProLeuLeuIleGlyAspValPheHisIleGlyLysTyrGluThrLysLys 316  
Db 273 AGAATTTCTGATATTTCTTCTGGGATGTTCGACAGTCAAGCAACCAAAATCTTCCGT 332  
QY 317 AsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgThrLeuAspArg 336  
Db 333 GCATTTGCTGCGAGGCGCGCTGAG-----GACCGC 365  
QY 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356  
Db 366 ACAATCGAATGATGATGCTGCTCAACCTGGCAATCGCAATCGCAAGATATTCGTGA 425  
QY 357 ThrGlnTyrArgPheAspGlyValValPhePheThrIleAspProLysThrAsnGluLeu 376  
Db 426 -----TTTCATTTGCGCGCATTTTCTCCCAAGATGAGCAATACATATGCGCAACAGACT 479  
QY 377 ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal 396  
Db 480 CACAGGAGAGAGATGAGATGATCTGTGAATGCAAGCAACCATCTTGAA----- 524  
QY 397 AsnMetGlyGluAlaTyrTrsLeuGlnAlaValAlaValAlaLeuSerAsnLeuIleAla 416  
Db 525 -----GACCTTGAAGCT 536  
QY 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGln 436  
Db 537 NACAG-----ATGATGATTCGAGACGAT-----GAGCGCATTCGATCAAG 581  
QY 437 Asn-----AspGlnVal-----SerPhe 442  
Db 582 AATACAGAGAAATGAATGATTTGTTTCATTTGATGATGATGATGATGATGATGATGATGAT 641  
QY 443 GluGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluPro 462  
Db 642 ACCTACGGGGCGT 692  
QY 463 ValIleGluThrValGluLeu-----ThrAspGlyIleLeu 474  
Db 693 TTGATGGAAGATGTTTCATTCAGACCAACAGTATGGAACAGTGAAGACACACCATATTCTT 752  
QY 475 MetAspIleSer----- 478  
Db 753 TTGCTGTCTCTGAGACATTTCACTCGGCCAAACCTTCAGACTTCTCCGTGAGCTCAA 812  
QY 479 -----ProIleGluThrSerAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAla 495  
Db 813 GGAAGGCTGCCATTCGCTGTGAATTTGAAGCCTTTGACCAAGAACAGACCTTTCTTGTATC 872



```

Db 914 AGAATACCTACCAATGGAGACCTTCCGACACCTCAATTAATGTTGATACCAAGTCC 973
Oy 380 GAspLysLeuProValLysArgGluLeu-----LeuGluGlnLeuLeuThrValAsnMe 398
Db 974 TACCAAGATG---ACAAAAGAAAGAGTTAAATGCTCAAGTCGACAGGGGATGGACCT 1030
Oy 398 tGlyLysValATyrAsnLeuGlnAlaValAlaLeuSerAsnAspLeuIleAlaThrAr 418
Db 1031 AGCAAGAGCCAGTGAAGTGGAG-----GTGAATAATGAATGCTGGCG----- 1073
Oy 418 gTyrPheAsnMetValAsnThrGluIleValPhe---ProGluArgGluGlnIleGlnAs 437
Db 1074 -----AATGTGGGAAAAAGAAATCTTGCAATACTGAGAAAGAACACACACAGA 1126
Oy 437 nAspGlnVal-----SerpheG 443
Db 1127 GGACACAGTGAAGAGCTGTGTGACATAGAGTATTCCTGCTGGTGAATAATACAGAGA 1186
Oy 443 uGlnSerSerSerSerArgThrGluPro-----AlaGlnValAs 456
Db 1187 CCAGAAATCTCTGAAGACACTGCCCCCTTCCAGAACCTTAGACAGTGTCTACCTATGA 1246
Oy 456 pGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAs 476
Db 1247 GGAACAGCTCAAGCCAAATCTTGAGAGCAGTCTCTCCCTGA----- 1292
Oy 476 pLaserProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAl 496
Db 1293 -----AACCACTACA 1303
Oy 496 aAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIle----- 513
Db 1304 GGTGAGTCAATGAGATGATGGTGCACCAAGTACAGACCAAGTCCCTTGAGCC 1363
Oy 514 -----AsnHisAspAspGlyValAlaAsnArgSerIle 523
Db 1364 ATCCAACTCTTGGAGTACTTGAATGTTGGAGAACACACAGAAATGTGGAGAGGACGC 1423
Oy 523 eLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspLys 543
Db 1424 ACTGACTCAGTTGAAGACAGCAGCAGCGCCGCTGCTGCTTGGGCGATAGTGA 1483
Oy 543 rGluAsnGlu-----ValIleAspLeuProGluArgThr----- 554
Db 1484 TCACACAGTTTATCATGATGACAAATGTATGTAGAGTCCCAAGAGTGAAGACAAG 1543
Oy 555 -----AlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysVal 571
Db 1544 CACAGAGCATAGTTTGAAGAAAGAAATTCACCAACAGGAAGCAGTGAAGCCCAAGAGGT 1603
Oy 571 lProLeuTyrValAlaPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGly 591
Db 1604 TCCAGCGCAC-----AGTACAGAGTAGTAGAGAT----- 1634
Oy 591 pGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAs 611
Db 1635 -----CACAAAC-----GAAAGACA 1648
Oy 611 pGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTy 631
Db 1649 GCGTGAAGAAACAGCA-----TTAAG 1669
Oy 631 rAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrG 651
Db 1670 GGACAGAAACCAATCAAGACAGAAAGATCTCTGCTTCCAGCAGGAATGAGGGCAATGT 1729
Oy 651 nGlnGluValAlaPheGlyHisSerThr 659
Db 1730 TCAGGAAGCGACAGGTCAAGTACA 1754

```

```

RESULT 27
BC018883          1876 bp      mRNA      linear      HTC 07-DEC-2001
LOCUS          BC018883

```

```

DEFINITION Homo sapiens, clone IMAGE:3617322, mRNA.
ACCESSION BC018883
VERSION BC018883.1 GI:17403016
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1876)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCMP/DRP
CDNA Library Preparation: Rubio Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Dierlich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Snyder,B., Stantrilop,S., Thomas,P.J.,
Tjongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
Zhang,L.-H. and Green,E.D.
FEATURES
source
1..1876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3617322"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIR_MGC_20"
/lab_host="DH10B-R"
/note="vector: pOMB7"
BASE COUNT 556 a 467 c 561 g 292 t
ORIGIN
Alignment Scores:
Pred. No.: 1 94 Length: 1876
Score: 102.50 Matches: 134
Percent Similarity: 31.5% Conservative: 90
Best Local Similarity: 18.90% Mismatches: 249
Query Match: 2.17% Indels: 237
Gaps: 32
US-09-914-168-2 (1-919) x BC018883 (1-1876)
Oy 43 ProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThr 62
Db 72 CCGTCCACCCCGAGCCGAGCTACGCCGGGCGGACCTGGCGCCGCTGTGGCGGGG 131
Oy 63 ProGluGlnIleGlnAlaArgLeuAsnAlaGlyLeuAsnAlaLysProGlnSerGln 82
Db 132 CCCGAGCCCGAGCGCCGAGCGGCTCCCGCGGCTGCTCTGCGCCGGAAGCGCGA 191
Oy 83 AlaLeuAspValValAsnHisAspAspGlnSerProIleSerThrGlyIleGlyGlnSer 102
Db 192 GCGTTCATCTTACCGCGGAGCTG-----GCTCCGTTCTCCGCGGACA---GAGCGCGG 239

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 12 Row: n Column: 11  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5174486  
 This clone has the following problem: no 5' EST match.

Oy 103 ProProLeuGlyLeuAspMetSerValIleGluGluThrProLeuSerLeuGlu 122  
 |||||  
 Db 240 CCCCCT-----GGCCCGCGCGGAGGGGGCTCCCGCG-GGGGTCCCGGAG 283  
 Oy 123 LeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnSplyrIleProGluTyr 142  
 |||||  
 Db 284 CATTTCCTCCG-----CCGGGTGGAGCGCGCGCGAGCCCGG 316  
 Oy 143 GlnGlyGlnGlnProAsnSerGluValValProProThrLeuGluProGluLysPro 162  
 |||||  
 Db 317 CAGAGATGACCGAGCCCGCGCGCGCG-----TCAAGCCG 349  
 Oy 163 GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnSplyrValAsnLysVal----- 180  
 |||||  
 Db 350 GGA-----GATCGACTGTTTGAGCCCGGAGAGCCGAGAGCGGCGGGA 391  
 Oy 181 -ProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGly 200  
 |||||  
 Db 392 AGCCCGGCTCGTCGCAAAAGCGCGCGCGCGCGGAGGCTCGGAGATCCGCATGAAGA 451  
 Oy 200 YSerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLalaLeuGluSpleIleTh 220  
 |||||  
 Db 452 GCTGGAGCCCGCAGACAGAGAGCTAGACAGACAGAA-----AAGATTTTAC 502  
 Oy 220 rGlnGlnSerAlaMetAspLeuAsnGly---SerIleProArgLeuArgGlnThrAlaLe 239  
 |||||  
 Db 503 TGAAGAGCGCTCGTAACTATGCGCGCGCTGTGCAGCCAGCGTGGCTGTGGGTGG 562  
 Oy 239 uValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle----- 254  
 |||||  
 Db 563 GACTCTCTCTGAGAGA---GGCAGCGGAGACCTCCATCCATCCATCCACACCGAGGCATC 619  
 Oy 255 -----IleArgAsnSerIleGlyGluValAlaSpleValIleIleHisAspLeuGly 271  
 |||||  
 Db 620 CATCAGGGAATCAAGACTCTAGACAGAACTTGAAGAG----- 659  
 Oy 271 uProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspSplys-- 290  
 |||||  
 Db 660 -----AAATATGAAGAGCTATGCTTCCCAATGCTCAGCTACATGAATAA 706  
 Oy 291 -----AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAs 305  
 |||||  
 Db 707 GACAACCTCATGTACCAGGTGTGATACCTTAAGATATGTGCCGAGCTTGAAGACA 766  
 Oy 305 rValPheHis---HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAl 324  
 |||||  
 Db 767 GCTGGCTGATCTAGCGCGGAGTACGAAGAAACAAAGAAATTGGAAGGGGAAACAA 826  
 Oy 324 aGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValSpleValIleLeuProAs 344  
 |||||  
 Db 827 CCCCCACACTACTG----- 842  
 Oy 344 rAsnThrAlaAspValSerLeuIleTyrIleAspThrGlyTyrArgPheAspGlyuVal 364  
 |||||  
 Db 843 -----CAATTCAGTTTGTGTAAGT 862  
 Oy 364 l-----ValPhePheThrI 369  
 Db 863 CAAGGAGCGCTGAGCAAGAGAGAAATGCTCGAACAACATGATATATCCTAATTC 922  
 Oy 369 eAspProLysThrAsnGlnLeuThrThrAsp-----Pr 380  
 |||||  
 Db 923 AGAAATAGTACCAATGACAGACTTCCGACACCTCAATATGTTGATACCAAGTCC 982  
 Oy 380 cAspLysLeuProValLysArgGluLeu-----LeuGlnGlnLeuLeuThrValAsnMe 398  
 |||||  
 Db 983 TACCAAGATC---ACAACAAGAGAGTTAAATGCCCTCAAGTCGACAGGAGTGGACCT 1039  
 Oy tGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnSpleuIleAlaThrAr 418  
 |||||  
 Db 1040 AGGAAGAGCCAGTGAAGTGGAG-----GTCAAAAATGAATCTGTGGCG----- 1082

Oy 418 gTyrPheAsnMetValAsnThrGluIleValPhe---ProGluArgGlnIleGlnAs 437  
 |||||  
 Db 1083 -----AATCGCGGGAAGAGAAATCTTGCAAAATATGTGAGAAAGAACACACAGACA 1135  
 Oy 437 nAspGlnVal-----SerPheGly 443  
 Db 1136 GAGACACAGTAAAGAGACTGTGTGCACATAGAGTATTCCTGTGTGTAGAGATACCGGAGA 1195  
 Oy 443 uGlnSerSerSerSerArgThrGluPro-----AlaGlnValAs 456  
 |||||  
 Db 1196 CCAAAATCTCTGTACACACTGTGCCATTCCTAGAACTTAGACAGTGTCTACTGTGA 1255  
 Oy 456 pGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAs 476  
 |||||  
 Db 1256 GGAACAGGTTCAAAAGCCAAATCTTGTAGACAGTCTCTCCGTGA----- 1301  
 Oy 476 pIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAl 496  
 |||||  
 Db 1302 -----AACACAGTACA 1312  
 Oy 496 aAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIle----- 513  
 |||||  
 Db 1313 GGTGAGTCAAAATGAGGTCTATGGGTGCACCATGATGACAGACAGAGACTCCCTTGAGCC 1372  
 Oy 514 -----AsnHisAspSplyValAlaAsnArgSerI 523  
 |||||  
 Db 1373 ATCCAACTGTGGAGTACTTAGATGTGGAGGCAACACAGCAAGATGTGGAGAGCGCAGC 1432  
 Oy 523 eLeuGlyArgIleSerAspAlaValAlaSerAlaValAlaArgAlaIleLeuProAspGluSe 543  
 |||||  
 Db 1433 AGTACTCAGAGTTGAAGAGAGAGCAGGACAGAGTGGCTGTCTTGTAGGGCATAGTGA 1492  
 Oy 543 rGluAsnGlu-----ValIleAspLeuProGluArgThr----- 554  
 |||||  
 Db 1493 TGACACAGTTTATCATGATGACAAATGTATGTAGAGGTCCCGCAAGATAGAGACAG 1552  
 Oy 555 -----AlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysVa 571  
 |||||  
 Db 1553 CACAGGCGCATAGTTTAAAGAAAGATTCACCAACAGAGAGCAGCTGAGCCAGGAGGT 1612  
 Oy 571 lProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTr 591  
 |||||  
 Db 1613 TCCAGCCGAC-----AGTACAGAAAGTAGTAGGAT----- 1643  
 Oy 591 pGlySerAspThrGlyThrArgLeuValIThrLysPheGlnHisAsnLeuIleAsnArgAs 611  
 |||||  
 Db 1644 -----CACAAC-----GAAGAGA 1657  
 Oy 611 pGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTy 631  
 |||||  
 Db 1658 GGGTGAAGAAACAGAG-----TTAAG 1678  
 Oy 631 rAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGly 651  
 |||||  
 Db 1679 GGACGAGAAACCAATTCAGACAGAGAGTCTGTCTTCACAGAGAGACTGAGGCAACTG 1738  
 Oy 651 nGlnGluValPheGlyHisSerThr 659  
 |||||  
 Db 1739 TCAAGAGCGACAGGTCCAGTAGTACA 1763  
 RESULT 28  
 BI086482 1014 bp mRNA linear EST 20-JUN-2001  
 LOCUS 602849752F1 NIH\_MGC\_10 Homo sapiens cDNA IMAGE:4991293 5',  
 DEFINITION  
 mRNA sequence.  
 ACCESSION BI086482 GI:14504812  
 VERSION BI086482.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1014)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN act:  
<http://image.llnl.gov>  
 Plate: L1008 row: h column: 14  
 High quality sequence stop: 463.

FEATURES  
 source location/Qualifiers  
 1..1014

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="491293"  
 /clone\_lib="NIH-MGC\_10"  
 /cell\_line="MGC36"  
 /lab\_host="DH10B"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."

BASE COUNT 213 a 325 c 299 g 177 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 0.749 Length: 1014  
 Score: 102.00 Matches: 60  
 Percent Similarity: 30.77% Conservative: 20  
 Best Local Similarity: 23.08% Mismatches: 87  
 Query Match: 2.16% Indels: 93  
 DB: 13 Gaps: 11

US-09-914-168-2 (1-919) x B1086482 (1-1014)

QY 573 LeuTYrValPheValAlaSerAspLysProArgASP-----GlyGln 586

DB 603 CTGCGGCTTCTGATAGACGGGTTCCCGGACAGAGGGCAATGACCGGGTCCAG 544

QY 587 TLeGlyLeuGlyTTPcLysSerAspThrGlyThrArgLeuValThrLysPheGlnHisAsn 606

DB 543 GTCAAGCTTGGGTG-----529

QY 607 LeuLLeAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlnAspLys 626

DB 528 -----CCCCAAGGTTTCAAGTTGCTGTCACAGCCAGA-----496

QY 627 GlyValLysLeuTYrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAla 646

DB 495 -----AACCTGGTTCTCCGCTGT 478

QY 647 ThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666

DB 477 CTTTGGGGGTCGCCGA-----GGGCATTCACTGGGGGAGAGAGAAATGGAAGTT 427

QY 667 ArgThrLeuGlnHisGluLysSerArgSerIleIleGlnAsnGlyLysTyrAsnArgThr 686

DB 426 -----TGGACCAAGTAA 415

QY 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaLarProGluThrTrp---705

DB 414 CTCCCTGTCAAGCCCGCTGCAAGCTTGGCCCTGCAAGGACGCTCGGAGCTTGGTTC 335

QY 706 -----GlnAspLeuPro-ValAspPheValAsnGlyLys 716

DB 354 AGCCCTCCCGTCGATCCACAGAATGCGGATGGTCCGGGCAACCTTCATATGTAATT 235

QY 716 sPSeRgGlnGluAlaLeuAlaGly-----ValAlaValHisLysThrYa 732

DB 294 GCCGGCCCAACCTGAGCTCTCGCAGCTGAAGCCGCGCGCGCCACCTTCGTGT 235

QY 732 LAlaAspAsnLeuValAsnProMetArgGly-----TyrArgGlnArgTyrSerLeu 750

DB 234 GGT-----CCGACCGTGGGACGCCACGATGGCCGAGGACCCG 190

QY 750 uValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLeu 770

DB 189 AGCGGGGACAGCGGGCTATTGCG-----GGCGCCTTGGCTTGGC 148

QY 770 rGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrG 790

DB 147 GGGCCATTACGTCGTGGGA-----TCTTACAGGCGCAGGACGATGTGAACCACTGGC 94

QY 790 yGlyIleGlnAlaGlyTyrIleTyrPserAspAsnHisValProTyrArgLeu 809

DB 93 CACGCCCGCTGCCATCTTGTGGAAGTGGGCTTCAAGACCATGCAATCGGCTG 36

RESULT 29  
 AK013402  
 LOCUS

DEFINITION  
 AK013402 2039 bp mRNA linear HTC 19-JAN-2002  
 Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length

enriched library, clone:2810473H05:homolog to  
 PHOSPHOCERYLGLYCOSAMINE PHOSPHATASE (PC 5.4.2.3) (PAGM)

(ACETYLGALUCOSAMINE PHOSPHOMUTASE) (N-ACETYLGALUCOSAMINE- PHOSPHATE  
 MUTASE), full insert sequence.

ACCESSION  
 AK013402.1 GI:12850738

VERSION  
 AK013402.1

KEYWORDS  
 HTC; CAP trapper.  
 Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA,  
 clone\_lib:RIKEN full-length enriched mouse cDNA library  
 clone:2810473H05.

ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
 Carninci, P., and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL  
 99279253

DB 10349636

REFERENCE  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL  
 20499374

DB 11042159

REFERENCE  
 3

TITLE  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multichannel sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
 20530913

DB 11076861

REFERENCE  
 4

AUTHORS  
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Konno, S., Yamana, T.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gasterland, T., Gissi, C., King, B., Koichiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
 Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,





OY	309	HisGlyLysTyrGluThrLysLysAsnLeu-----IleGluAsnAlaSerLacIuHis	326
Db	717	TCAGTTAAAGTTGATGACTGCCAACGGCATAGGGCCCTTGAAACTAGACAAATGGAACAC	776
OY	327	GlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThr	346
Db	777	---TACTTC-----TCCGGGGCCCTGCGTT-----	800
OY	347	AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPhe-----	361
Db	801	-----CTGCTGTTAATGATGGGACCCCAAGGGCGCTCAATCACCTTGCGGT	848
OY	361	-----	361
Db	849	GCTGACTTTGTCAAAAGTCAACAGAAACCCCAAGGCAATTGAATGAGTCCGGTGAG	908
OY	362	-----AspGluValAlaPhePheThrIleAspProLys	372
Db	909	ACATGCTGCTCTTCGTGGATGGGATGGGACAGAGATCGTGATTACTACGTGATGAGAT	968
OY	373	ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGln	392
Db	969	GCTCACTTTCATCTCATAGATGAGACAGAAAGAGCGACTTAATTAAGCGTTTCCCTTAA	1028
OY	393	LeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsn	412
Db	1029	GAGCTACTCTTGTGAGATTGGAAGAAGTGTGAACCTCGGAGCTGACAGCAGCATATGCA	1088
OY	413	AspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArg	432
Db	1089	AATGGAAGTTCACACGGGAGTACTGTGAAGAAGTTATGAAGGTACCTGTGATTGACCAAA	1148
OY	433	GluGlnIleGlnAsnAspGln-----ValSerPheGlu	443
Db	1149	ACTGCTGTGTTAAACATTGTGCATCAACAAGGCTCAAGATTGTGACATCGAGTTTATTGTGA	1208
OY	444	GlnSerSerSerArgThrGlnProAlaGlnValAspGlnSerThrLeuGlnProVal	463
Db	1209	GCGAACGGGCATGGAACAGCACTGTTCACTGAGACAGTTGAAGTGAAGATAAAGACACTA	1268
OY	464	IleGluThrValGluLeuThrAspGly-----IleLeuMetAsp	476
Db	1269	GCCCAA-----GAACGTAGACGATGGGAAGAAAGACAGCCAGGACGGCTGCGAGCATC	1322
OY	477	IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla	496
Db	1333	ATCGACTTGTTCACACGAGCACTGAGAGATGCTATTTCGACATGTGGTGATTGAGGCT	1382
OY	497	AlaLysAlaArgHisLeuTyrAspMetProAspArgValLeuAlaIleAsnHisAsp	516
Db	1383	ATCCGAGGCTCTG-----AAGGGCTGACTGTACAGCAGCTGG	1418
OY	517	AspGlyVal-----AsnArgSerIleLeuGlyArgIleSerAsp	529
Db	1419	GATGCTATTATGTGCATCTTCTTAACACAGCAACTAAAGTTAAAGTTCGCGAGGAGACA	1478
OY	530	-----AlaValSerAlaValAlaAlaArgAlaIleLeuProAspGlnSerIleAsnGlu	546
Db	1479	GTTATTAGCACCAAGGATGCTAGAGACAAAGAGTCACTCCACCAAGACTCCAAAGAGCA	1538
OY	547	ValIleAspLeuProGluIuArgThrAlaLeuAla	557
Db	1539	ATCATGACTGCTGTCAGAGAAATACACACTTGGC	1571
RESULT 30			
LOCUS	A2927049	513 bp	DNA
DEFINITION	476.dist16c05.sl Saccharomyces castellii NR1_Y-12630 Saccharomyces		
ACCESSION	A2927049		
VERSION	A2927049.1	GI:13497952	
KEYWORDS	GSS.		

SOURCE	Saccharomyces castellii.					
ORGANISM	Saccharomyces castellii.					
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.					
AUTHORS	1 (bases 1 to 513) Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M. Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis Unpublished (2001)					
TITLE						
JOURNAL	Contact: Johnstone M Department of Genetics Washington University Medical School Box 8332, 4566 Scott Ave., St. Louis, MO 63110, USA Tel: 314 362 2735 Fax: 314 362 7855 Email: mjgenetics.wustl.edu					
COMMENT	Class: random plasmid subclone. Location/Qualifiers					
FEATURES	source 1..513 "/organism='Saccharomyces castellii'" "/strain='NRRL Y-12630 (CBS 4309)'" '/db_xref='taxon:27288' '/clone='476.01s16c05.s1' '/clone_lib='Saccharomyces castellii NRRL Y-12630' '/note='Random genomic sequence'					
BASE COUNT	181 a	102 c	90 g	140 t		
ORIGIN						
Alignment Scores:						
Pred. No.:	0.257	Length:	513			
Score:	101.50	Matches:	47			
Percent Similarity:	44.51%	Conservative:	30			
Best Local Similarity:	27.17%	Mismatches:	51			
Query Match:	2.15%	Indels:	45			
DB:	17	Gaps:	11			
US-09-914-168-2 (1-919) x A2927049 (1-513)						
OY	372	LysThrAsnGlnLeuThrTrpProAspLysIleuProValLysArgGluLeuGlu	391			
Db	5	AAAATAATATAAACACCAAAACAAGT-----GAAGTCATTGAA	46			
OY	392	GInLeuLeuThrVal---AsnMetGlyGluAlaTyrrAsnLeuGlnAlaValargalau	410			
Db	47	AAGTAGCTACTATTATGAAACAACGGT-----AAAGCTTTA	82			
OY	411	SerAsnAsp-----LeuIleAlathArGtyrPheAsnMetValasphrculuile	427			
Db	83	AcGAtGTgATGATTGATTCTCAGACCACGACTACTTTATCGAATTCGCCACC-----	136			
OY	428	ValPheProGluAuArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer	447			
Db	137	----CCCTGCACCATCGAAGAATGGAATTTTCGGAGAATCGTTCACCACGCGTTGAAA	190			
OY	448	SerArGthrGluproAlaGlnValaspGlu-----SerThrLeuclupro	462			
Db	191	TCGAATTAACAAGATCTTAATTGAGATCAAGATCTATATCCAGTCGTACTCGCG	250			
OY	463	VallieGuthrValGluLeuThraspgLyIleLeuMet-----AspIleSer	478			
Db	251	ATTTATCAAAGGTTCATCCGAAAAACAATTATATGTGGTGATACGAATCAGAGT	310			
OY	479	ProIlegUpheSerAlaSerAsn-----LeuIleOlnaspLysLeuasnleu	494			
Db	311	ATTATGATGATTTTAGGGGATTAATTCAGAATTAATAAGTCCTTGATTAAGCTCACCTG	370			
OY	495	ValalaialySalatrgHisIeuleuyraSp-----MetProasp	507			
Db	371	ACCAATTCACGACTCTCCACTTTATGATTAACCTTCGATGCACACCGAAATTAATAAT	430			
OY	508	-----AsparValleualaiIleasnHIsaspDspgly	518			

```

Db      431   TCCCGCAACAGCATTCCTGTCA---AACCATCATCAGGA 466

RESULT 31
BC028159          2340 bp      mRNA       linear      HTC 01-MAY-2002
LOCUS           BC028159
DEFINITION      Homo sapiens, similar to KIAA0076 gene product, clone
IMAGE:5240443, mRNA.
ACCESSION      BC028159
VERSION        BC028159.1  GI:20380216
KEYWORDS
SOURCE         Homo sapiens.
ORGANISM       Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2340)
AUTHORS        Strausberg,R.
TITLE           Direct Submission
JOURNAL         Submitted (08-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK         NIH-MGC project URL: http://mgc.nci.nih.gov
COMMENT         Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgt.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho S.-L., Karlins,E., Latic,P., Legaspi,R., Maduro,Q.L.,
Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
Pearson,R., Stanchilpop,S., Thomas,P.J., Touchman,J.W., Tsugeon,C.,
Vogt,L.L., Walker,M.A., Wetherby,K.D., Wiglins,L., Young,A.,
Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAX Plate: 62 Row: 1 Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction
This clone has the following problem: frame shifted.

FEATURES             location/Qualifiers
     source            1..2340
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5240443"
                     /tissue-type="Brain, fetal, whole pooled"
                     /clone_lib="NIH_MGC_121"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-Sport6"
BASE COUNT          519 a      631 c      721 g      469 t
ORIGIN
Alignment Scores:
Pred. No.:          4.35              Length:          2340
Score:              101.00            Matches:          109
Percent Similarity: 32.90%            Conservative:     70
Best Local Similarity: 20.04%          Mismatches:      199
Query Match:        2.14%              Indels:          166
DB:                 11                  Gaps:            27

US-09-914-168-2 (1-919) x BC028159 (1-2340)
Qy      84 leuaspaValaIasnPh easpaSpai nserProIIeSerArgIlleglyClucInIs erPro 103
||||| :|||:|
DB      858 CTGGATCGACTGCAAT-----AGCATGTCGAGAGCTGGAGAGTACCAAGCTCC 908
||||| :|||:|

```

QY	104	Pro-----	LeuGlyLeuAspMetSerVal	111
Db	909	CCATGTCGCACAGAGAGAAAACCCGGGAGCAGCGGAACTGGATTCGACGTGGCTGTG		968
QY	112	IleGluIuThrThrProLeuSerLeuGluGluPheAlaGlnGluSerThrgLmet	131	
Db	969	GGCAACCTCACTCTGTAGCTTTGCGGAGACATGGGCTGGAACTCAAGCAAG	1028	
QY	132	GlyIleAsnProAsnAspTryIleProGluTryGlnGluGlnProAsnSerGluVal	151	
Db	1029	GGCATGTCACCT-----	CCCGGCGCAACCGGCTGCATC	1061
QY	152	ValValPro-----	ProThrLeu-----	GluProGlu
Db	1062	TTTCAGCCCTPACATTTTCAGGCCCCACCCCTTTACTTCGCCACCATTTGTCCACACCCCGACA	1121	
QY	161	LysProGlyLeuIleLeuAsnArgLeuTryAlaArgLeuPheAsnAspGlyValAsnLysVal	180	
Db	1122	AGACAAAGGTGGGTC-----		1136
QY	181	ProArgLeuLysAlaLysPheTryGlnSerSerGlnSerGlyGluThrSerAlaIleGly	200	
Db	1137	-----	TTCCGCCAGCCCTCGATTTCTCAGCCGTAGTGCGCTATGGA	1178
QY	201	SerSerHisGlnLysThr-----	GluProTryAlaAsnLysAlaAlaLeuGluAspIle	219
Db	1179	GAATATGTGCAGACACATCGACAGCCAGGAGATGGCGTGGC-----	ATGCTGATGATATAT	1235
QY	220	ThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu	239	
Db	1236	GAGGAGATCACTGCT-----	GGGACAGAGGCGCACTCCGCGACGC-----	1277
QY	240	ValAlaAlaArgAlaValAlaGlyTryAspIleAspLeuSerIleIleArgAsnSerIle	259	
Db	1278	-----	AAACAGCGCATT	1289
QY	260	GlyGluValAspValIleIleHisAspLeuGlyGluProValTryIleAspTryAlaArg	279	
Db	1290	CCCCCTGGCGAGGTCTTGTGGACATCGACAGCGCGCATTTACTGGGTGCACATGGACATG	1349	
QY	280	ValGluValArgGly-----	GluGlyAlaAspAspIleAlaPheThrThrValAlaAsp	297
Db	1350	CTGAGATCTCTGGGCGCCTGACGAACCCACAGAGATAAAGCTTCACAGCATGTGGACAG	1409	
QY	298	GluVal--	ProLeuLeuIleGlyAspValPheHisHisGlyLysTryGluThrLysLys	316
Db	1410	GGGCGAGGGCTACTGCTGTGGCGACACGATTT-----		1442
QY	317	AsnLeuIleGluAsnAlaSerAlaGlnHisGlyTryPheAspGlyArgTyrLeuAspArg	336	
Db	1443	-----	CCCTCTGGGAGTGGATTCCTATGGATGGGCTGTAC-----	1478
QY	337	SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTryAspThrGly	356	
Db	1479	CCTTTGGCGCTACCTCCAGCCCGAAGCTCAGAGAATGAGAGATGGAGATAT-----	CTG	1532
QY	357	ThrGlnTryArgPheAspGlyValValPhePhe-----	ThrIleAsp-----	370
Db	1533	ACCACGGCTGATAGTGGGAGCTGCTTTTCTTTATCAAAAAGTTGGACTGTGTGACAG	1592	
QY	371	--	ProLysThrAsnGlnLeuThrThrAspProAsp-----	381
Db	1593	CAGCCAAATTTTCCGAATCTTTTGGAGAACCTGGATGACACCCTGGGTGMAAAGCCCTA	1652	
QY	382	--	LysLeuProValLysArgGluLeuGluGlnLeuLeuThrValAsnMetGlyGlu	400
Db	1653	GGTGATGATCTGTGTCCCGGAAATGAGCGAGAGTCTGTGAGAGTT-----	CTCAGTAGT	1709
QY	401	AlaTryAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTryPhe	420	
Db	1710	CGATTTGGGGCGACACTGCTCAATGATGCTGTCTCACTCCAGATCTTACACCAAGTATGGG	1769	

QY	421	AsmValAsn	-----	ThgIuIleValPheProGlu	431
Db	1770	CTGCTGCTATGAACAACAGACGCTGCTTACTTCACGAATCACTCTGATCCCAAT	111		

**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**  
**REFERENCE**  
**AUTHORS**

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861

4  
 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Ozawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Ikazaki, Y., Gojobori, T., Bono, H., Kusunawa, T., Saito, R., Kodate, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Mochizuki, L., Mashima, J., Mazzatelli, J., Momtaz, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y.

**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**  
**REFERENCE**  
**AUTHORS**

Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851

5 (bases 1 to 2936)  
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Aikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imochani, K., Ishii, Y., Itoh, M., Izawa, M., Kasekawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M., Koyama, S., Kuhlra, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numata, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, H., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**  
**REFERENCE**  
**AUTHORS**

Direct Submission  
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGACAGATCCAGACCTCTTTTCTTTTCTTTTATN 3'], cDNA was prepared by using triethanolamine-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGACATTCCTGAGTTAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SDR.

**FEATURES**  
**SOURCE**

1. .2936  
 location/Qualifiers  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM.DB:2610044J23"





```

Db 679 CTGGAGAGCCGAGAAACGTGATTAAGCTGGAGCCAGGAGTCATCTCGMGGGGGCGAG 738
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 333 r-----AlaGluHisGlyTrpPheAspGlyArg----- 332
      |||      |||      |||      |||      |||      |||      |||      |||
Db 729 TCCCATCTCATCTGATCTACCTCCATCGACAGACCGAGCTTGAGAGTGCTTGAGCT 798
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 333 -----TrrPheAspArgSer-----ValAspValIleLeuProAspAs 345
      |||      |||      |||      |||      |||      |||      |||      |||
Db 799 CCGCATCCAGCTGGCGGCGACATCAGGACGACCATGTGCCATCATCTCTGCGGCA 858
      |||      |||      |||      |||      |||      |||      |||      |||

Oy 345 nThrAlaAspValSer 350
      |||      |||      |||      |||      |||      |||      |||      |||
Db 859 CAGGCGACAGCTTGCC 874

RESULT 34
LOCUS   BF256095/c
DEFINITION BF256095 800 bp mRNA linear EST 22-OCT-2001
            HVSMET0008L17f Hordeum vulgare seedling root EST library HVCDNA0007
            (Etiolated and unstressed) Hordeum vulgare cDNA clone
            HVSMET0008L17f, mRNA sequence.

ACCESSION BF256095
VERSION   BF256095.2 GI:31118036
KEYWORDS
SOURCE    Hordeum vulgare.
           Hordeum vulgare.
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
           1 (bases 1 to 800)
           |||      |||      |||      |||      |||      |||      |||      |||
REFERENCE 1 (bases 1 to 800)
AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Fritsch,D., Yu
           ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
           ,R.D., Oates,R. and Main,D.
           Development of a genetically and physically anchored EST resource
           for barley genomics: Morex unstressed seedling root cDNA library
           Unpublished (2001)
           On Nov 16, 2000 this sequence version replaced gi:11185208.

JOURNAL   Contact: Wing RA
           Clemson University Genomics Institute
           100 Jordan Hall, Clemson, SC 29634, USA
           Tel: 864 656 7288
           Fax: 864 656 4293
           Email: rwing@clemson.edu
           Total hg bases = 287
           Seq primer: AATTAACTCTCACTAAAGG
           High quality sequence stop: 644.

FEATURES
         source
           1..800
           /organism="Hordeum vulgare"
           /cultivar="Morex"
           /db_xref="taxon:4513"
           /clone="HVSMET0008L17f"
           /clone_lib="Hordeum vulgare seedling root EST library
           HVCDNA0007 (Etiolated and unstressed)"
           /tissue_type="Seedling root"
           /lab_host="TJCl21"
           /note="Vector: lambdaZAP; Site:1: EcoRI; Site:2: XhoI;
           Seeds were surface sterilized then germinated under axenic
           conditions in the dark at room temperature on filter paper
           with water, nystatin and cefotaxime in covered
           crystallization dishes. Five-day old seedling roots were
           then harvested, total RNA was prepared, poly(A) RNA was
           purified, one primary unamplified cDNA library was made,
           and 1 million pfu were in vivo excised to give plus/crypt
           SK(-) cDNA phagemids. These steps were performed in the TU
           Close laboratory at the University of California,
           Riverside (Choi, Close, Fenton). Phagemids were plated and
           picked at the Clemson University Genomics Institute (CUGI)
           (Begum, Palmer, Fritsch, Atkins and Wing). Plasmid DNA
           preparations, DNA sequencing and sequence analysis were
           performed at CUGI (Wing, Yu, Fritsch, Henry, Simmons, Oates
           , Rambo, Main). The sequence has been trimmed to remove

```

```

vector sequence and contains a minimum of 100 bases of
pred. value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders
see Close T.J., Wing R., Kleinhofs A., Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/g99pages/bgn/31/cover.html)"

BASE COUNT 91 a 217 c 352 g 140 t
ORIGIN

Alignment Scores:
Pred. No.: 0.858 Length: 800
Score: 100.00 Matches: 70
Percent Similarity: 33.22% Conservative: 32
Best Local Similarity: 22.80% Mismatches: 140
Query Match: 2.12% Indels: 65
DB: 12 Gaps: 9

US-09-914-168-2 (1-919) x BF256095 (1-800)
Oy 14 ProValAlaLeuAlaLeuAlaLeuProLeuMetThr-SerGlnAlaLeuAlaGlnGlnAs 33
      |||      |||      |||      |||      |||      |||      |||      |||
Db. 789 CCGGGGCGCAACCGCGCATACAGCAAAACACCACCCAGCCGACGACGCTGGCGACGA 730
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 33 nAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAlaIleAsnGlnAla 53
      |||      |||      |||      |||      |||      |||      |||      |||
Db 729 CATCCCGCC----- 720
Oy 53 salagIAsnProProValLeuLeuThrProGlnGlnIleGlnAlaArgLeuAsnAla 73
      |||      |||      |||      |||      |||      |||      |||      |||
Db 719 -----CCCCCAATGATACCCCCCCCCACACAA---AACGGCACTACAGCTCCG 673
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 73 agILeuAsnAlaIleProGlnSerGlnAlaLeuAspValaIleAsnPhAspGlnSe 93
      |||      |||      |||      |||      |||      |||      |||      |||
Db 672 CGGTCTCACCAGCCGCCAGATCCCGACCATCTC-----GAGTC 634
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 93 rProIleSerArgIleGlyLeuIleSerProProLeuGlyLeuAspMetSerValIleG 113
      |||      |||      |||      |||      |||      |||      |||      |||
Db 633 CCGCGCTCTCGACAGCCAGGATACCACCCCTCCAGCCAGCAAAACAGCATCCCG 574
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 113 uGluThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGly 133
      |||      |||      |||      |||      |||      |||      |||      |||
Db 573 ACCCACCCGCGCGTCTC----- 555
Oy 133 eAsnProAsnAspTrpIleProGluTrpGlnGlnIleProAsnSerGluValaVala 153
      |||      |||      |||      |||      |||      |||      |||      |||
Db 554 -CGCCCGCGGCCCATCCCGCACACAGGCCCCCAGCCAGCCCGCCAGGCCAGCAGC 496
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 153 lProProThrLeuGluProGluIleProGluIleIleAsnArgLeuIleAlaArgLeu 173
      |||      |||      |||      |||      |||      |||      |||      |||
Db 495 CCCCCCTGTGGCCCGCCCTCGGCCGCCACCTCCCATGGCCAGCCACAGGCCAC 436
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 173 e-AsnAspGlyValaIleAsnIleValProArgLeuIleAsnIle-LysPheTrpGlnSerGln 192
      |||      |||      |||      |||      |||      |||      |||      |||
Db 435 CCGAGCCCGGACGCGCCAGGTCCTCCAGTCTCCCGGACCGGTCCTCATGTCACCCG 376
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 193 SerGlyLeuThrSerAlaIleGlySerSerHisGlnIleThrGluProTyAlaAsnIle 212
      |||      |||      |||      |||      |||      |||      |||      |||
Db 375 CAGGTGCGGGGCCAGCGCGTGGGGAGAGTCCGACAGCACACACCA---GCCGGTGTG 319
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 213 LysAlaIleLeuGluAspIleThrGlnGlnIleSerAlaMetLysPheAsnGlnSerIlePro 232
      |||      |||      |||      |||      |||      |||      |||      |||
Db 318 GCGCTTGCGCCCTCCCAAGGTCACCGCGGATGTACTTCCACTCCAGCGAGGTGTTC 259
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 233 ArgLeuArgGlnThrAlaLeuValaIleAlaArgAlaValaGlyTrpTrpAspIleAspLeu 252
      |||      |||      |||      |||      |||      |||      |||      |||
Db 258 CAGCTT-----GATGTCGGGTGCGGAGCGGA-----GACGATCTG 220
      |||      |||      |||      |||      |||      |||      |||      |||
Oy 253 SerIleIleArgAsnSerIleGlyGluValaIleIle----- 266
      |||      |||      |||      |||      |||      |||      |||      |||
Db 219 TCGGTGTCAATACAGAAAGCGACGAGAGGAGGACGTCCTCTCGTGGCGGAGGCGCG 160

```

OY	267	---	HisAspleuGlygluIproValTyrIleAspTyrArgAlaValGluValArgGlyGlu	285
Db	159	GGTGGTGACCCCGGTGGAGAGGTCCACAGAGCGGAGCTCTCCAGATCCGGAAG	100	
OY	286	-----	GlyAlaAspAspLysAlaPheTrpThrValAlaAspGluValPro	300
			:::	
Db	99	CGCGAACrTGGCGCTGGAGAGCTGAGACGACrCTCrGGACAACAAAGTGGCGCTGTAG	40	
OY	301	LeuLeuIleGlyAsp	305	
Db	39	GCGGCGGAGGGGAC	25	
RESULT 35				
AL523772				
LOCUS				
DEFINITION	AL523772	1041 bp	mRNA	linear
	AL523772	LTL.NFL003.NBC3	Homo sapiens	CDNA clone CS0DC0031M01 5
ACCESSION				
VERSION	AL523772			
KEYWORDS	AL523772.1	GI:12787265		
SOURCE		EST.		
ORGANISM		human.		
		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE				
AUTHORS		1 (bases 1 to 1041)		
TITLE		Li,W.B., Gruber,C., Jesssee,J. and Polayes,D.		
JOURNAL		Full-length CDNA libraries and normalization		
COMMENT		Unpublished (2001)		
		Contact: Genoscope		
		Genoscope - Centre National de Sequencage		
		Bp 191 91006 EVRI cedex - France		
		Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES		Location/Qualifiers		

source	1. 1041		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="LRI_NFL003_NBC3"		
	/sex="male"		
	/tissue_type="neuroblastoma cells"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <a href="http://fulllength.invitrogen.com">http://fulllength.invitrogen.com</a> "		
BASE COUNT	217 a	283 c	280 g 259 t
ORIGIN	2 others		
Alignment Scores:			
Pred. No.:	1.37	Length:	1041
Score:	100.00	Matches:	67
Percent Similarity:	37.25%	Conservative:	47
Best Local Similarity:	21.90%	Mismatches:	96
Query Match:	2.12%	Indels:	96
DB:	9	Gaps:	16

[illegible]

OY	558	AsnArgLysThrProAla-----AspValTyrGlnSerLysLysValPro	572		
Db	332	CCTGTGAAGTCGCCCGCCTTGATTCAGTGCCTGGGACATGTAGACTCTCTCTGGCTCC	391		
OY	573	LeuTyrValPheValAlaSerAspLysProArGAspGlyInIleGlyLeuGlyTrrpely	592		
Db	392	AAGTAACCTCCC-----	403		
OY	593	SerAspThrGlyThrArgLeuValThrlLysPheGluHisAsnLeuIleAsnArgAspGly	612		
Db	404	-----CATACCTTGTTTGGATCACCA	427		
OY	613	---TyrGlnAlaGlyAlaGlyLeuArgLysSerGluAspLysLysGlyValLysLeu---	630		
Db	428	GTTTACAGTCTGTGGTTCGAAGAACAATCATATATTCAAAGAATCTTCAGATYGAG	487		
OY	631	-----TyrAlaThrLysProLeuSerHisProLeuAsnAspGln---LeuArg	645		
Db	488	ATTTTTGGATGTTTTGCTCTGTACCGAATTTAAACCAGGACAGTAATACCAAGGCCATTGGC	547		
OY	646	AlaThrLeuGlyTyrGln-----GlnGluValPheGlyHisSerThrAsnGlyPhe	662		
Db	548	ACAACCTGCCCATCAGATCTGCACACTGAGGAATTCATCATATATCCCCCT-----	598		
OY	663	AspLysSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyGly	682		
Db	599	-----GATTTCGAAGCTGCCAAGTTTGGTTGGCCAACATGAGC	637		
OY	683	TTPAsnArgThrTyrSerLeuArgTyrArgLeu-----AspLysLeuLys	697		
Db	638	AAGACAGCACTCAACCGCGTGTCTTAAAGCTGTGTGTGCCAAGGAGACCATGGCCAT	697		
OY	698	ThrGlnAlaPro-----ProGluThrTrpGlnAspLeuPro---	709		
Db	698	GGGCTCATCCCTTTATCGTCGAGATCCGGGACCCGGAAGACCTTCCATGCTGTGA	757		
OY	710	----ValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyValAla	727		
Db	758	GTGATGTGTGGCGACATGAGAAAAAACCTCGGGCAGAACGGCTGTGATTAATGTTTCGCG	817		
OY	728	Val--HisLys-----Thr	731		
Db	818	ATGTTCCACAGAGTCAGAGTTCCTCGCACAGACCTTTCGAACCGGATGGAGACGTCACCC	877		
OY	732	ValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluVal	751		
Db	878	CCCGAGGGCACCATTATGTACAGCCCTTTTAAAGCAGCTCAGACGACGCTTTGGAGCGTCCCG	937		
OY	752	Gly-----SerSerGlyLeuValSerAspAlaAsnMetAla---IleAlaArgAlaGly	768		
Db	938	GGGAGCCGTGTCTCCGGGCGGTCTTCATCTGTAGACTGGGACCTTCCTTAACATAAGCTGGC	997		
OY	769	IleSerGlyValTyrSer	774		
Db	998	GTGGCATCGCTCTTCCT	1015		
RESULT_36	AKO19509	3499 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AKO19509				
DEFINITION	Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632435C11:CARBOXYPEPTIDASE X2, full insert sequence.				
ACCESSION	AKO19509				
VERSION	AKO19509.1				
KEYWORDS	HTC; CAP trapper;				
SOURCE	Mus musculus (Strain:C57BL/6J) 0 day neonate skin cDNA to mRNA, clone:4632435C11.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1	Carninci,P. and Hayashizaki,Y.			
AUTHORS					





Dh 386 CCCAAGAGCCATCAGCCCAAGAGGCTCCCAAGAGGAGACTTACTTGCCAGAGACG 445  
Qy 71 AsnAlaIaIaGlyLeuAsnAlaLys-----ProGlnSerGlnAlaLeuAspVal 87  
Dh 446 CCTCCACAGCGTAAATAATGACAACAGAAAGGAGAAAGCAAGAAATTGTGAGAAAGCT 505  
Qy 88 AsnPheAspArgLys-----ProIleSerArg-----IleGlyLysIleSerProPro 104  
Dh 506 GCCAGTGAAGACCATGCTGCTGCTCATGAGATGTCAGAGAGAGATTGCCACCT 565  
Qy 105 LeuGlyLeuAspMetSerValIleGlyLysIleThrProLeuSerLeuGlyLysLeuPhe 124  
Dh 566 CTTCGGCTCG-----GAACATTAAATACACAGACTTCCACGCTGCAT 607  
Qy 125 AlaGlnLysLys-----ThrGlnMet 131  
Dh 608 GGCTCCACATCGAAGCGTTATGGCCGAGGAGCCACGGGGGAGACTCAACATCCAGGCA 667  
Qy 132 GlyIleAsnProAsnAspTyrIleProGlyTyrGlnGlyLysIleProAsnSerGlyVal 151  
Dh 668 GGCATTAAATGAATAATGACTTT-----TACGATGGGGCTTGGTGTGCTGTAGGAAC 718  
Qy 152 ValValProProThrLeuGlnProGlyLysProGlyLeuIleLysArgLeuTyrAlaArg 171  
Dh 719 GACTTCATCATGATGCATCGAAGTGATGCCGGCGCTGACCAAG-----TTCACAGGG 772  
Qy 172 LeuPheAsnAspGlyValAsnLys-----ValProArgLeuLysAla 185  
Dh 773 GTTCATTACCCAGGAAGAAAGACTCTCTGCTGAGTACGATGGGTGACATCTTAAGATC 832  
Qy 186 LysPheTyrGlnSerSerGlnSerGlyLysIleThrSerAlaIleGlySer----- 201  
Dh 833 ATGGTGAGCAATGACAGCCACACATGCTGTTACTGTGAGAAATGAGATCTGGCGACATGATA 892  
Qy 202 -----SerHisGlnLysThrGluProTyrAlaAsn-----IleLys 213  
Dh 893 TTGTGAAGGAACAGTGAAGAGAGATTCTGTGCTCAATGAGCTGCCAGTCCCATGTGTG 952  
Qy 214 AlaAlaLeuGlnAspIleThrGlnLysIleSerAlaMetAspLeuAsnGlySerIle----- 231  
Dh 953 GCCCGCTACATTCGCATTAACCTCAGTCTGCTGTAT-----AACGGAGCATCTGCATG 1009  
Qy 231 ----- 231  
Dh 1010 AGGATGAGATCTTGGGCTGCCACCTGCCGATCTTAATACATTATACCGACCTAAT 1069  
Qy 232 -----ProArgLeuArgGln 236  
Dh 1070 GAGATGACCAACCAAGATACCTGGATTTTAAGCACCAACTATTAAGAAATGGCCAG 1129  
Qy 237 ThrAlaLeuVal----- 240  
Dh 1130 TTGATGAAGTGTCAATGAATGTGCCCAATATTACAGAGATTATACAACTTGGCAAA 1189  
Qy 241 AlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGly 260  
Dh 1190 ACCCAACCAAGGCGCTGAATGTATGCGAGAGATCTCT-----AACCATCTCTGG 1240  
Qy 261 GluValAspValIleIleHisAspLeuGlyLysProValTyr-----IleAspTyr 277  
Dh 1241 GAA-----CATGAAGTTGGTGAAGCCGAGATTCCACTACATCGAGGGGC 1285  
Qy 278 ArgAlaValGluValArgGlyLys----- 285  
Dh 1286 CACGGCAATGAGATTCTGGAGAGAACTGCTGCTGCTGCATCTTCTCTCCAG 1345  
Qy 286 -----GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIle 303  
Dh 1346 GAATTACTGGGCGAGAACCAAGCATGTCCTGCTGTGAGAGACTCGAATCCACATT 1405  
Qy 304 GlyAspValPheHisIleSclLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSer 323  
Dh 1406 CTACCTCCCTCAATTCCTATGCTATGAG-----AAGGCTATGAAGAGGTTCC 1456

Qy 324 AlaGlnHisGlyTyrPheAspGlyArgTyrProLeuAspArgSerValAspValIleLeuPro 343  
Dh 1457 GATTGGAGAGCGTGGTCCCTGGAGAGCTTGACCCATGATGGCATGCATATC-----AAC 1510  
Qy 344 AspAsnThrAlaAspVal-----SerLeuIleTyrAspThrGlyThrLysIleThrArgPheAsp 362  
Dh 1511 AACCACTTCGGAATTTAAATCTGCTGCTGGAGGACAGAGCAG----- 1558  
Qy 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrAspProAspLys 382  
Dh 1559 -----CAGATCCCAACGAAAG 1576  
Qy 383 LeuProValLysArg-----GluLeuLeuGlnLysLeuThrValAsnMetGlyGlnAla 401  
Dh 1577 GTCCCAACACACTACATTCGCATCCCTGAGTGTTCTGTCTGAATGCAACAGTGGCC 1636  
Qy 402 TyrAsnLeuGlnAlaValArgAla----- 409  
Dh 1637 ACAGAGACCAAGACCCCTCATCGCTGATGAGAGATCCCGTTTGTCTGGAGGCAAC 1696  
Qy 410 -----LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleVal 428  
Dh 1697 CTACAGGGGGGTGAGCTGTGCTGCGCATACCCCTATGACATGCTGCGGTCCCTGTGAAG 1756  
Qy 429 PheProGluArgGlnIleGlnAsnAspGlnVal-----SerPheGlnGln 444  
Dh 1757 ACCCAGAGACACACCCCAACACTGATGATCATGTCTCCGCTGGCTGGGTATCTTCTAC 1816  
Qy 445 SerSerSerSerArg-----ThrGluProAla 453  
Dh 1817 GCCTCCACTCACCGCCTCATGACAGATGCCAGAGGCGAGTGTGCCACAGGAATTTT 1876  
Qy 454 GlnValAspGlySerThrLeuGlnProValIleGlyThrValGluLeuThrAspGlyIle 473  
Dh 1877 CACAAGAGAGAGGCGACCCGTCATGGGGCT-----TCTGGAGACACAGTGGTGAAGT 1930  
Qy 474 LeuMetAspIleSerProIleGlnPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn 493  
Dh 1931 CTAAAGCATTTTCAG-----TACCTCCATACAAAGCTGTTCGATGCTGCATCTAC 1981  
Qy 494 LeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp-----Arg 509  
Dh 1982 GTGGCGCTGTGATTAATACCAACGAGACGAGACCTCCGAGGAAGGAGAAATAACCCG 2041  
Qy 510 ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAsp 529  
Dh 2042 GAGTCTGTGATTTGTTCATGAGCAGGTTTCATCGAGCATCAAGGCAATAGTACAGAT 2101  
Qy 530 AlaVal-----SerAlaValAlaArgAlaIleLeuProAspGlnSerGlnLysGlnValIle 548  
Dh 2102 TTACAGGGAAGAGGATTTCAAATGCTGTGCATCTGTGGAAGGTATTAACCATGACATC 2161  
Qy 549 AspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSer 568  
Dh 2162 -----CGGACACCC-----AGCGATGGGAGATTACTGCGCTCA 2194  
Qy 569 LysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyIleIleGly 588  
Dh 2195 CTGAACCCGTGGGCAATATGTGTGCACACGCAAGCGCGAA----- 2233  
Qy 589 LeuGlyTyrPheLysAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIle 608  
Dh 2234 ---GGCTTATACCTTCCACCAAGAACTGATGTT----- 2266  
Qy 609 AsnArgAspGlyTyrGlnIleGlyAla-----GluLeuArg 620  
Dh 2267 -----GGCTATATATGGAGCTACTCGGTGACTTCAACCTCAACAAAGCAAC 2317  
Qy 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisPro 639  
Dh 2318 CTGGCTAGGATTAAGAAATTTATGAGACATTTTGGAGACAGACGCTCTCAGCGCTACCC 2374

RESULT	37
CNS013XL	
LOCUS	CNS013XL 969 bp DNA linear GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN10105 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL103443
VERSIONS	ALI03443.1 GI:5615054
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster. Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epmydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 969)
REFERENCE	Genoscope. Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequence BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  Web : www.genoscope.cns.fr ) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CERH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES	Location/Qualifiers  1..969 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACN10105" /cclone_1lb="DrosBAC" /plasmid="pbelOBAC11" /note="end : sp6"
BASE COUNT	288 a      243 c      222 g      184 t      32 others
ORIGIN	
Alignment Scores:	
Pred. No. :	1..38          Length:          969
Score:	99.50          Matches:         36
Percent Similarity:	45.77%        Conservative:    29
Best Local Similarity:	25.35%        Mismatches:     54
Query Match:	2.10%         Indels:          23
DB:	17             Gaps:          7
US-09-914-168-2 (1-919) x CNS013XL (1-969)	
Oy	23 LeuMeThSerGIAlaleuaIagInGlnAsnProAlasniLeileasnHlsVal 42 :::                                  ::::
Db	110 GTRATGGCAACCTCAGCGMTTGGAGCACAG-----GAGCTGGCTAACCCGGTG 157
Oy	43 ProAHlIsaprHaIalIeasnGlAlalyAlaylasn-----ProPval 59                ::::            ::::
Db	158 -----CATTTTCACCACACAGCAGCAGTAGTGCAGTGTCTCAGTCCCACAAT 211
Oy	60 LeuDeuThrProCIugInleGIalnAargLeuasnaIalaGLyleuasnaIaLysPro 79 :::::                                 ::            ::::
Db	212 GTGGTTAACTCTTAAGAACAACAGCAGCCGTCACACACAGACTCCACAAACGAGCT 271
Oy	80 GlInsErAlaleuaspaValaIsnpheaspasplInsErPrOllieserArg----- 97 :::::          ::            ::            ::::
Db	272 ACAAMTGGCGCTACTACGAATGCCACACGCCGAGACCTGCTCGGTGTCACAGACTGT 331
Oy	98 -----IlleglygUGInsErProProleucglyLeuaspmetserValIIegluGluThr 115 :::            ::::               :::::            ::
Db	332 CCCCTACTTCGCTACCCACCTCTCCGAACCTGTTCAGCAGCCTATAATCATCAATA 391
Oy	116 ThrProLenSerleuGIngluleuPheaIaglIngusErThrgImetGcylIleasnPro 135                ::::            ::               ::
Db	392 ACTAGGCCGACGTGTGTCACACTGCTATGCCCAAAAAAGTCACA-----CTACCA 439

Oy	136	Asna\$prYr1leproGluTyGInclYgluGInprOa\$nsSergluVal-----Val	152
Db	440	TCA-----CCAGTTTACGAGGCGGAAGTCAAAAGTTAGTTGAGAGCATT	490
Oy	153	ValPro 154	
Db	491	GTGCCA 496	
RESULT 38			
LOCUS	AY108441	1514 bp	mRNA
DEFINITION	Zea mays	PC0083558	mRNA sequence.
ACCESSION	AY108441		
VERSION	AY108441.1	GI:21211519	
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC		
AUTHORS	clade: Panicoideae; Andropogoneae; Zea.		
TITLE	1 (bases 1 to 1514)		
JOURNAL	Hainey, C.F., Dolan, M., Miao, G. H., Vogel, J.M., Whitesitt, M.S.,		
REFERENCE	Arthur, L.W., Hanfey, M., Morgante, M. and Tingey, S.V.		
AUTHORS	Maize Mapping Project/DuPont Consensus Sequences for Design of		
JOURNAL	Overgo Probes		
TITLE	Unpublished (2002)		
REFERENCE	2 (bases 1 to 1514)		
AUTHORS	CoE, E.C.		
JOURNAL	Direct Submission		
TITLE	Submitted (25-APR-2002) Maize Mapping project, University of		
JOURNAL	Missouri, Columbia, MO 65211, USA		
FEATURES	Location/Qualifiers		
source	1..1514		
	/organism="Zea mays"		
	/db_xref="maizedb:634674"		
	/db_xref="taxon:4577"		
	/clone="PC0083558"		
	/clone_1db="Maize Mapping Project/DuPont Consensus		
	Library		
	/note="this sequence is part of a project of EST		
	assemblies resulting from the application of public		
	contigs to seed DuPont contigs; this resource was		
	assembled by DuPont as part of a collaboration for the		
	overgo addressing of BACs in conjunction with the Maize		
	Mapping Project"		
BASE COUNT	452 a 323 c 371 g 368 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.05	Length:	1514
Score:	99.50	Matches:	74
Percent Similarity:	35.48%	Conservative:	64
Best Local Similarity:	19.02%	Mismatches:	146
Query Match:	2.10%	Indels:	105
DB:	11	Gaps:	15
US-09-914-168-2 (1-919) x AY108441 (1-1514)			
Oy	193	SerglygluThrSerAlaIleGlySerSerHisGlnlys-----	205
Db	65	TCAAGCGACGAGATGGGATGGCGCGCGCTCGAGCGAGGAGGAGCCCTCATCTCTCC	124
Oy	206	---ThrgluProTyrrAlaAsnIleLeYsAlAlaAlaLeuGluAspIleThrgIngluSerAla	224
Db	125	TTCACCGCTCCCAACCCGCTCTGTGGCGCGCGCTTCCCTCTGGCGTACGAGCATCTCCGCC	184
Oy	225	MetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAla	244
Db	185	GTCTCGACACGAGATGGCTCCA-----	208
Oy	245	ValIlyTyrrTyrrAspLeuSerIleIleArgsn-----SerIleGlyglu	261
Db	209	CTGTGTCGACGCGCTATTTCACACTCAATCTGTGAGGACCGAGTCAAGACGTGTCAGGAT	268

```

OY 262 ValAspValIleIleHisAspLeuGluProValTyrIleAspTyrArgAlaGlu 281
    ::::: ||| ::::: ||||| :::::
Db 269 ATCCAGACATATCACAAGGCTATGAGATGGTTGCAGCATCAAGCTCCGGAGTTGAG 328

OY 282 ValArgGlyGluGlyAlaAspAsp-----LysAlaPheThrThrValAlaAspGluVal 299
    ||| ||| ::::: ||||| :::::
Db 329 ACAGAGACTGAAATTCACAGCTGCTGTGGCAACCGCTGCACCCCTCTTGGCGATGCT 388

OY 300 ProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIle 319
    |||
Db 389 CCT-----AGTCTTGACGTCAAGAAATGTGATT 418

OY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAsp 339
    ::::: ||| |||
Db 419 GTGGCCATTACCTCTCATTAAGGACCTCTGTGCTGT----- 454

OY 340 ValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyr 359
    ||| |||
Db 455 -----ATTAAATTCACATCAGTGAAA--- 475

OY 360 ArgPheAspGluValValPhePheThrIleAspProLysThrAsnGluThrThrAsp 379
    :::::
Db 476 -----GTTACGAGGCGCCCTTCACAAAATTCAGATCTGCT 508

OY 380 ProAspLys----- 508
    |||
Db 509 CCAGAAAAAAGAAACCAAGTATGTTATATTAGAGAAAGGCAAGTTGACGCTGTGGCT 568

OY 385 ValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGlnTyrAsnLeu 404
    ::::: |||
Db 569 GACTCAAGAGACAGCATTTGAAATGACTGCAGCACTGCAGAAAGAACCTATCACTAC 628

OY 405 GlnAlaValArgAlaLeuSerAsnAspLeuIle-----AlaThrArg 418
    ||| |||
Db 629 ACACAGCTTCCCTGCTCGCGGATGACATTTTCAAAAATGTGCAATATGATGCTCGAGG 688

OY 419 ---TyrPheAsnMetValAsnThrGluIleValPhe---ProGluArgGluGlnIleGln 436
    ||||| ||| ||| |||
Db 689 GTTATTTTCAACAAATTCATCTGTCATTCATTCCTTAAGCCACAAATGTAACAAATACT 748

OY 437 AsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp 456
    ::::: ||| :::::
Db 749 TCCCCCGAGCTTGCA---GAAAAAGATAGCAAGATGCTGGGAAGATGGCTGACCTAG 805

OY 457 GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp 476
    ||||| |||
Db 806 TCTTATGAG-----ATTGAAGCGCGGAGACAAATCAGAGATTTTGCAGAT 853

OY 477 IleSerProIleGluPheSer-----AlaSerAsn 486
    ::::: ||||| |||
Db 854 CTAGCTGAGTTCCAGTTTCTGTGTCATGATCAATGCTGCCCTAGAGATGATGACGT 913

OY 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
    ::::: ||||| |||||
Db 914 GAGCTTGGAGCCCGTATGCTCTGCCATGAGACGCTTACGAAATCTGCGCATATGCTT 973

OY 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
    ||| ||| |||
Db 974 GATCGCTCACTCTCACTTACAAACAGACAGC-----CAACATTCATCACCACAGAG 1027

OY 527 IleSerAspAlaValSerAlaValAla 535
    ::::: |||
Db 1028 CTTACTGACATTTATCTGTGTCATCG 1054

RESULT 39
AK002434 1562 bp mRNA linear HMC 19-JAN-2002
LOCUS AK002434
DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched
library, clone:061009009:DNA segment, Chr 15, Wayne State
University 77, expressed, full insert sequence.
ACCESSION AK002434
VERSION AK002434.1 GI:12832416

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
HMC; CAP trapper.
Mus musculus (strain:C57BL/6J) adult male kidney cDNA to mRNA,
clone.lib:RIKEN full-length enriched mouse cDNA library
clone:061009009.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2049374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system--384-Format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arai, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gasterland, T., Gissi, C., King, B., Kochiya, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nakado, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staudt, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., But, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5 (bases 1 to 1562)
Adachi, J., Aizawa, K., Akahira, S., Akiyama, T., Aono, H., Arai, A.,
Arai, K., Baldarelli, R., Bono, H., Brownstein, M., But, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyama, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, S.,
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shikata, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.

```



Db 1115 -----CAGAGGCCAGAGATGTGGCCGCAACTTCGAGAGACAG----- 1159  
 Qy 747 TyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArg 766  
 Db 1160 -----GAGCTCTCCATACCCCAAGCTGGCGCTGATGAGATGCGCACCTACCGC 1210  
 Qy 767 AlaGlyIleSerGlyValTyrSer-----PheGlyAsp----- 777  
 Db 1211 AAGCTGCTGGAGGGGAGAGAGAGCCGCTTCTGTGAGAGACGCAATGCGACTGTGAATATC 1270  
 Qy 778 -----AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAla 794  
 Db 1271 TCTGTGTGAATTCACATGGGGCAATGTCGACAGCTTATCTCTCGAGAACCATGGGA 1330  
 Qy 795 GlyTyrIleTyrPserAspAsnPhenAlaValProTyrArgLeuAlaPheMetAlaGly 814  
 Db 1331 AGCAATGCTCTGAGC-----TTCAAGTGGG 1354  
 Qy 815 GlyAspGlnSerIleArgGlyTyrAla 823  
 Db 1355 GGACCCGGAGCGCTCAGGGCCTATTC 1381  
 RESULT 40  
 W29713 1022 bp mRNA linear EST 11-SEP-1996  
 W29713/c  
 LOCUS  
 DEFINITION  
 IMAGE:338008.5' similar to gb:V00714 mouse gene for alpha-globin (mouse);, mRNA sequence.  
 W29713  
 VERSION  
 KEYWORDS  
 W29713.1 GI:1309862  
 EST.  
 SOURCE  
 ORGANISM  
 Mus musculus  
 house mouse.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1022)  
 Maier,M., Hillier,T., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisell,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.  
 The Washu-HMI Mouse EST Project  
 Unpublished (1996)  
 Contact: Maria M/Mouse EST Project  
 Washu-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.wustl.edu  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
 MGI:219408  
 Seq primer: ETPRimer  
 High quality sequence stop: 350.  
 Location/Qualifiers  
 1. 1022  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone\_image="338008"  
 /clone\_id="Soares mouse p33MF19.5"  
 /dev\_stage="19.5 dpc total fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Vector: pT73 (pharmacia) with a modified polylinker Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento

Soares and M. Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University).  
 BASE COUNT 223 a 225 c 307 g 266 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.75 Length: 1022  
 Score: 99.00 Matches: 80  
 Percent Similarity: 38.74% Conservative: 49  
 Best Local Similarity: 24.02% Mismatches: 122  
 Query Match: 2.09% Indels: 83  
 DB: 14 Gaps: 19

US-09-914-168-2 (1-919) x W29713 (1-1022)

Qy 30 AlaGlnGlnAsnAspProAlaAsnIleIleAsnHis-ValProAla-----HisAs 46  
 Db 923 AGCAAGAGAACACACCATTCATCCCGCATATCCTCCCGCCAGCATATTCATCA 864  
 Qy 46 pThrAlaIleAsnGlnAlaValAlaGlyAsnProProValLeuLeuThrProGluGlnI 66  
 Db 863 CCTTTTTCGATACACCAAA-----AACCCCTTTTAAACCCACCAAAAGGCGCG 810  
 Qy 66 eGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaValSerProGlnSerGlnAlaLeuAspVa 86  
 Db 809 ACAATCCCGCAT-----CCACGCGCAAAATTTCTTG----- 779  
 Qy 86 lValAsnPhenAspAspGlnSerProIleSerArgIleGlyGlnGlnSerProProLeuGl 106  
 Db 778 -----GAAAGGCCAATC--CGTGTGG----- 758  
 Qy 106 lLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluLeuPheAlaGl 126  
 Db 757 -----CTTTTCCGATATCTCCCTCCACTCTCCAAAAGGGGATTTCTGTCA 711  
 Qy 126 nGluSerThrGluMetGlyIleAsnProAsnAspTyr-----IleProGluTyrGlnI 144  
 Db 710 ACGCAGATTAATTCGGG---AACCCACAGATTTCACCAATGACCCAGATTGAA 654  
 Qy 144 yGluGlnProAsnSerGlyValAlaValProProThrLeuGluProGluLys----- 161  
 Db 653 ACAACCGCCAAATCCCAAACTAAATTAACCTTACCTTAAGAGAGAGATTAACCTTCCG 594  
 Qy 162 -ProGlyLeuIleLysArgLeuTyr----- 169  
 Db 593 CCCATGTTATTTTTCATCTCTCTCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 534  
 Qy 170 -AlaArgLeuPheAsnAspGlyValAlaValProAlaGlyLeuLysAlaLysPheTyrGl 189  
 Db 533 TCAGGCTTTATTCAAAGCATGATGTACAGTTGCAACGAGAGAAACAGGATGCCA 474  
 Qy 189 nSerSerGlnSerGlyGluThrSer-----AlaIleGlySerSerHisGln----- 204  
 Db 473 GAATGCAGACCCCGCAGAACGACGCTTAACGGTACTTGGAGGTACACGCTGCTCACAG 414  
 Qy 205 -----LysThrGluProTyrAlaAsnIleLysAlaAla-----Le 216  
 Db 413 AGCAAGAAATTTATTCAGAGAGCATTCACCGCGGGGGAATGCGAGGTGGCT 354  
 Qy 216 uGluAspIleThrGlnGlnUser-----AlaMetAspLeuAsnGlySerIleProAr 233  
 Db 353 ACCCAAGGTACCCAGAGGAGCTGTGCTGAGAGCTTGAAGTTCAGCGATCCACACAGCAG 294  
 Qy 233 gLeu---ArgGlnThrAlaLeuValAlaAlaIleArgAlaValGlyTyrTyrAspIleAsp 252  
 Db 293 CTGTGGGATCAGATGCTCTGAGACAGACAGGAGCGGAGGTCTATCGAG-GTGGC 235  
 Qy 252 uSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluPr 272  
 Db 234 CTGCAGCATTTGGCCAGAGCATCGGCGACCTTTCGCGAGACCCCTTGAGCGGAGAGC 175  
 Qy 272 oVal---TyrIleAspTyrArgAlaValAlaValArgGlyGluGlyAlaAsp----- 288

```
Db 174 CGTGGCTTACATCAAGTAGGAGAGTAGTCTTGTGGTGGGAAGCTAGCAACATCC 115
QY 289 -----AspLysAlaPheThrValAlaAspGluValPro-----Le 301
Db 114 TTTCAGGGCTTCAGCTCCATATTCAGCACCATGGCCACCAATCTTCCCCAGGCGCCT 55
QY 301 uLeuIleGlyAspValPhe-----HisHisGly 310
Db 54 TGATGTGCTTTTGTCTCTCCCGACAGACACCATGTGT 18
```

Search completed: May 6, 2003, 01:45:34  
Job time : 2095 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 6, 2003, 01:32:51 ; Search time 669 Seconds  
(without alignments)  
6472.581 Million cell updates/sec

Title: US-09-914-168-2  
Perfect score: 4727  
Sequence: 1 MSKPVIFANRSEMPVALAAY.....TGCKEENGRKILHFIQGPFP 919

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6113114 seqs, 2355906770 residues  
Total number of hits satisfying chosen parameters: 12226228

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09914168.r/unat\_28042003\_151617\_6869/app\_query.fasta.1.1095  
-DB=Pending\_Patents\_NA\_New -QMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -DOOPCT=0  
-DOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09914168.@CGN.1.1.465.@runat\_28042003\_151617\_6869 -NCPU=6 -ICPU=3  
-NO\_XUPXY -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-NARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending\_Patents\_NA\_New.\*

1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

No.	Score	Match	Query	Length	DB	ID	Description
1	620.5	13.1	1818	8	US-10-419-128-14784		Sequence 14784, A
2	620.5	13.1	1818	9	US-10-366-683-14784		Sequence 14784, A
3	500.5	10.6	1860	8	US-10-417-886-3279		Sequence 3279, Ap
4	436.5	9.2	1737	6	US-09-913-101-1		Sequence 1, Appl
5	427.5	9.0	1731	6	US-09-913-101-3		Sequence 3, Appl
6	405	8.6	4497	8	US-10-419-128-14663		Sequence 14663, A
7	405	8.6	4497	9	US-10-366-683-14663		Sequence 14663, A

C	8	362	7.7	726	8	US-10-419-128-14966	Sequence 14966, A
C	9	362	7.7	726	9	US-10-366-683-14966	Sequence 14966, A
C	10	353	7.5	1416	9	US-10-144-771-21979	Sequence 21979, A
C	11	351	7.4	1365	10	US-60-453-134-81	Sequence 81, Appl
C	12	319	6.7	944	8	US-10-399-416-9	Sequence 9, Appl
C	13	265.5	5.6	2379	9	US-10-148-534-6	Sequence 6, Appl
C	14	259	5.5	2391	9	US-10-320-800-5	Sequence 5, Appl
C	15	259	5.5	2394	9	US-10-148-534-2	Sequence 2, Appl
C	16	259	5.5	2394	9	US-10-148-534-10	Sequence 10, Appl
C	17	238.5	5.0	6617	7	US-09-689-065A-1	Sequence 1, Appl
C	18	238.5	5.0	6617	7	US-09-689-065B-1	Sequence 1, Appl
C	19	206.5	4.4	3012	6	US-09-988-067B-7	Sequence 7, Appl
C	20	200	4.2	2778	8	US-10-335-977-157	Sequence 157, App
C	21	186	3.9	2466	8	US-10-419-128-4442	Sequence 4442, Ap
C	22	186	3.9	2466	8	US-10-366-683-4442	Sequence 4442, Ap
C	23	174.5	3.7	516	8	US-10-419-128-14662	Sequence 14662, A
C	24	174.5	3.7	516	8	US-10-366-683-14662	Sequence 14662, A
C	25	162.5	3.4	1530	8	US-10-419-128-4165	Sequence 4165, Ap
C	26	162.5	3.4	1530	8	US-10-366-683-4165	Sequence 4165, Ap
C	27	160.5	3.4	2163	9	US-10-282-122A-33708	Sequence 33708, A
C	28	158	3.3	1812	9	US-10-010-160-15	Sequence 15, Appl
C	29	157	3.3	1890	8	US-10-369-493-42248	Sequence 42248, A
C	30	154	3.3	2349	10	US-60-453-134-135	Sequence 135, App
C	31	154	3.3	3111	9	US-10-282-122A-27975	Sequence 27975, A
C	32	152	3.2	2183	8	US-10-369-493-44808	Sequence 44808, A
C	33	150	3.2	2763	8	US-10-282-122A-33378	Sequence 33378, A
C	34	148.5	3.1	4802	9	US-10-263-929-30	Sequence 30, Appl
C	35	148.5	3.1	4802	9	US-10-267-502-24	Sequence 24, Appl
C	36	146	3.1	6609	9	US-10-172-502-1	Sequence 1, Appl
C	37	145.5	3.1	7542	9	US-10-320-800-39	Sequence 39, Appl
C	38	145.5	3.1	7545	9	US-10-282-122A-29937	Sequence 29937, A
C	39	145.5	3.1	2272325	9	US-10-018-470A-1	Sequence 1, Appl
C	40	145.5	3.0	4473	9	US-10-282-122A-29386	Sequence 29386, A
C	41	143.5	3.0	3920	9	US-10-282-122A-14980	Sequence 14980, A
C	42	142.5	3.0	4293	9	US-10-282-122A-42267	Sequence 42267, A
C	43	142.5	3.0	1752	8	US-10-369-493-23985	Sequence 23985, A
C	44	142	3.0	1752	8	US-10-282-122A-37677	Sequence 37677, A
C	45	141.5	3.0	5631	9		

## ALIGNMENTS

RESULT 1  
US-10-419-128-14784  
; Sequence 14784, Application US/10419128  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/10/419,128  
; CURRENT FILING DATE: 2003-04-21  
; PRIOR APPLICATION NUMBER: US/09/252,991  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ. ID NOS: 33142  
; SEQ ID NO 14784  
; LENGTH: 1818  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-10-419-128-14784

## Alignment Scores:

Pred. No.: 8.39e-41  
Score: 620.50  
Percent Similarity: 40.61%  
Best Local Similarity: 24.45%  
Query Match: 13.13%  
DB: 8  
Length: 1818  
Matches: 177  
Conservative: 117  
Mismatch: 225  
Indels: 205  
Gaps: 15

```

US-09-914-168-2 (1-919) x US-10-419-128-14784 (1-1818)
OY 210 AlaAsnIleLysAlaIleLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGly 229
    |||||
Db 199 GCCAATATATCGAGCGCTAGCTGGTACGCGGAGCGGAC-----GAGGCG 246
OY 230 SerIleProArgLeuArgGlnThrAlaLeuVal-----AlaAlaArgAlaVal 245
    |||||
Db 247 GCGTTGCAACGTTTCCGTCGCAATGCCGAGCGCGGAGGAAAGCGGCGCCAGCGGCTC 306
OY 246 GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle 265
    |||||
Db 307 GCGTACTTCCAG-----GCCCAATCATCGACGCGAG 336
OY 266 IleHisAsp-----LeuGlyGluProValTyr 274
    |||||
Db 337 GTGAAGACGCGCAAGCGCGGCAAGCTCAACGTAAGTCCCGGCGAGCGCTGGCG 396
OY 275 IleAspTyrArgAlaValAlaGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThr 294
    |||||
Db 397 CTGCGCCAGGTCAACATCCAGGTGCTCGGAGCGCGGAGCGCTGGAGAGCTTCGCGTTG 456
OY 295 ValAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThr 314
    |||||
Db 457 CCGTCGCGCGAG---CAACTGAAGCGGCGCGGAGCTGAACGAGGTCTACGAGGAC 513
OY 315 LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeu 334
    |||||
Db 514 GCCAAGCGCGGTATCCACACGACGCGCTCGCGTACGCTTCCAGCGCGCTTCGAGC 573
OY 335 AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerIleTyrAsp 354
    |||||
Db 574 ACCGAGCGCCCTGACATGATCCCGCGCGCGCATCGCCGACATCGACTGCTTACGAC 633
OY 355 ThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsn 374
    |||||
Db 634 AGCGCGCAGCGCTACCTTCGCGCAGGTGAGCTTC----- 669
OY 375 GlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlnLeuLeu 394
    |||||
Db 670 -----GACGCGCATCCATC---ATCGAGAGAAGACGTGTCGCGCATGCTG 714
OY 395 ThrValAsnMetGlyAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu 414
    |||||
Db 715 CCGTTCAAGCGCGCAACCTTAGTACTCGAAGTATCCGCAACATCAACACAGACTG 774
OY 415 IleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGln 434
    |||||
Db 775 CAATCCAGCGCTACTTC----- 792
OY 435 IleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGln 454
    |||||
Db 792 ----- 792
OY 455 ValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeu 474
    |||||
Db 793 -----GAGGAGAGTGGCG 804
OY 475 MetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeu 494
    |||||
Db 805 GTGAGCGCGCGCGGAGCCGAGCTCAGGCC----- 834
OY 495 ValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsn 514
    |||||
Db 834 ----- 834
OY 515 HisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaVal 534
    |||||
Db 835 -----GACCGCGCGCGCGAGCGGATC----- 855
OY 535 AlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThr 554
    |||||
Db 855 ----- 855

```

---

```

OY 555 AlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyr 574
    |||||
Db 856 -----CCGGTACCG 864
OY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAsp 594
    |||||
Db 865 CTCGCGCTGGAGAGCGCGGACCGGACCGATGGCGGTCGCGGCTCGGTTCGACCGCAC 924
OY 595 ThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGln 614
    |||||
Db 925 GTGCGGCGCGCGGAGGTTCACCTGAGCCCGCATGGGTGGATGCCAGAGACATATGC 984
OY 615 AlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuThrAlaThrLys 634
    |||||
Db 985 CTGGATTCGAGTGGAAATCTCCGACACCGAGGAGAAATCGCGGCGCTGTACGAGATT 1044
OY 635 ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluVal 654
    |||||
Db 1045 CCCCTCATCCGCGCGCTACCGAGACAGCTGCGCTTCAACAGCGGCTTCCAGATTGCA-- 1101
OY 655 PheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGlyIleSer 674
    |||||
Db 1102 -----GACSTAGTCGACCGGACGAGCAAGCTGACCTTGGCGCGGAA 1146
OY 675 ArgSerIleIleGlnAsnGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAsp 694
    |||||
Db 1147 TGGCAGACGAGAGCGCGCGATGGCTGACGCGGTGCTTCGGAATCGATCGCGGAG 1206
OY 695 LysLeuLysThrGlnAlaProProGluThrTyrProLysProValAspIleValAsn 714
    |||||
Db 1207 GAGTACAAAGCTG----- 1218
OY 715 GlyLysProSer-----GlnGluAlaLeuLeuAlaGlyAlaValAlaHisLysThrVal 732
    |||||
Db 1219 GGGCAGACATCCGCGGCTGAGACACTTCGTATGCGGGATGCTGACTGCTGAGAG 1278
OY 733 AlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValAlaGly 752
    |||||
Db 1279 ACCGACAAAGGTGAGAGCCGACCATGGTACCGCTGAGATTCAACGCAAGGGGCGG 1338
OY 753 SerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyVal 772
    |||||
Db 1339 AAGGAGGCGCTGTGGCGCGAGCTGCTCATATGTCAGTGGCCAAAGGGCGCTG 1398
OY 773 TyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIle 792
    |||||
Db 1399 ACCAGCTTGC CGCG-----GCCATCGCTGCTGCGCGCTG 1437
OY 793 GlnAlaGlyTyrIleThrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePhe 812
    |||||
Db 1438 CAGGTAGCGGGAATCCGCCAACAGCTACAGTCAATCCGCGCGCTGCTTTC 1497
OY 813 AlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp 832
    |||||
Db 1498 GCTGGCGGAGCCAGAGCGCTGCTGCTGACTGACCGGACCTGTGCGCGGAGGAATCC 1557
OY 833 LysGlyTyrLeuThrGlyGlyGlnValIleuAlaValAlaGlyThrAlaGluThrAsnTyrGlu 852
    |||||
Db 1558 GATGGGAGAACAGATCGCGCGCGCTACATGATCCCGGAGCGTGAAGATCAATATCCG 1617
OY 853 PheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGly 872
    |||||
Db 1618 CTGGCGAGCGCTGCGCGCTGCGACCTTCGTCAGACAGGCGGACGCTTCACATCGCTG 1677
OY 873 PheThrAsnAspThrLysIleGlyAlaGlyValAlaGlyValArgTyrPheSerProValGly 892
    |||||
Db 1678 GACTTCCCTCGATCAAGACCGCGGTGCTGCGCGCTGCGCTGCGCTGCGCGCTGCGG 1737
OY 893 GlnValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHis 912
    |||||
Db 1738 CCGTTGGCGCTGACCTGGCCCATGCGCTCGACGACGAGCGGCT--TTCCGCTGACAC 1794

```





```

Db 1279 ACCGACACAAAGGTGCGACCCAGCCATGGCTACCGCTGCATTCAACGTCAAGGGGGCG 1338
Oy 753 SerSerGlyLeuValSerAspAlaSerMetAlaIleAlaIrrAlaGlyIleSerGlyVal 772
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1339 AAGGAAGGCTGCTGGCCGACGCCGACGCTCTTCATGCTGACGCCGCAAGGGGCTG 1398
Oy 773 TyrSerPheGlyAspAsnAlaTyrGlySerAsnAlaGlnHisGlnMetThrGlyIle 792
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1399 ACCAGCTTCCCGCGC-----GGCCATCGCGCTGCGCGCGCGCTG 1437
Oy 793 GlnAlaGlyTyrTrpIleTrpSerAspAsnPheAsnHisValProTyrAlaGlyLeuPhePhe 812
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1438 CAGGTAGGCGGAAATCGCCACCAACGACTACAGTCCCGCTGCGCTGCGCTTCTTC 1497
Oy 813 AlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp 832
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1498 GCTGGGGGCGACAGAGCGTGGGTGATGACGATACCGGACGCGTGGCGGAGAAATTC 1557
Oy 833 LysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGlu 852
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1558 GATGGCGACAAAGATCGCGCGCGCTACATGATCGCGCGCGCTGAGTATCAATATCCG 1617
Oy 853 PheMetLysAspLeuArgLeuAlaValAlaPheGlyAspIleGlyAsnAlaTyrAspLysGly 872
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1618 CTGGCGGAGCGCTGGCGCTGGCGGACCTGCTGACACAGGCGCAAGCCTTCAACTGCTG 1677
Oy 873 PheThrAsnAspThrLysIleGlyAlaGlyValAlaGlyValaTrrAlaSerProValGly 892
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1678 GACTTCCCGCTGATCAAGACCGGGGTGCGCTGCGCTGCGCTGGGTCTGCGCGGTGCG 1737
Oy 893 GlnValaArgValaAspAlaIleAlaThrGlyValaGlyGlnGluGlnLysProIleLysLeuHis 912
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1738 CCCTTGGCGCTGACCTGGCCCATGCGCTGACGACGACGCGCGGT--TTCCGCTTGAC 1794
Oy 913 PhePheIleGly 916
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1795 TTCTCCATGCGG 1806

RESULT 3
US-10-417-886-3279 : Application US/10417886
: Sequence 3279, Application US/10417886
: GENERAL INFORMATION:
: APPLICANT: Keith G. Weinstock et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
: TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.135
: CURRENT APPLICATION NUMBER: US/10/417,886
: PRIOR FILING DATE: 2003-04-17
: PRIOR APPLICATION NUMBER: US/09/252,691C
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/094,145
: PRIOR FILING DATE: 1998-07-24
: PRIOR APPLICATION NUMBER: US 60/074,787
: NUMBER OF SEQ ID NOS: 11326
: SEQ ID NO 3279
: LENGTH: 1860
: TYPE: DNA
: ORGANISM: Enterobacter cloacae
US-10-417-886-3279

Alignment Scores:
Pred. No.: 5,65e-31 Length: 1860
Score: 500.50 Matches: 183
Percent Similarity: 38.09% Conservative: 100
Best Local Similarity: 24.63% Mismatches: 253
Query Match: 10.59% Indels: 207
DB: 8 Gaps: 19

US-09-914-168-2 (1-919) x US-10-417-886-3279 (1-1860)
Oy 196 ThrSerAlaIleGlySerSerHisGlnLysThrGlnProTrrAlaAsnIleLysAlaIle 215
    ||||| ||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 169 ACCAGCGGATTCGCAAGCGC-----GCGAATGCTCGTTTGCAG 207
Oy 216 LeuGluAspIleThr-----GlnLysSerAlaMetAsp 226
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 208 GTTGAGGGGTTATTCGGGGGCGTGGAAAAAACGTGCGTGGCAAGCTTTTACTATTCAG 267
Oy 227 LeuAsnGlySerIleProAlaGlyArgGlnThrAlaLeuVal----- 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 268 AGCGATGAGGTCAAGCGCGATTCGCGCTTTGCGCGCGCGTGGATGACGCATTCGGAA 327
Oy 241 AlaAlaArgAlaValGlyTyrTyrAsp-----IleAspLeuSerIleIleArgAsnSer 258
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 328 GCGTTAAAGCGTGGGGGTGCTAGCAACCCACCATGATTCGATTATTCAGTTCGCGC 387
Oy 259 IleGlyGluValAspValIleIleHisAspLeu-----GlyLysProValTyrIleAsp 276
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 388 GCAAAGAGACGTGAGGTGATTCGCGCGCGTCCGCGCGGCGCAACCGCTGATTCGCGC 447
Oy 277 TyrArgAlaValGlnValaArgGlyGlyAlaAspAspLysAlaPheThrValAla 296
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 448 GGCACGACGTGGTGGTGGCGCGCGCGCGCGCACGCGATGCTGACTGATCTGCTC 507
Oy 297 AspGluValProLeuLeuIleGlyAspValPheHisIleGlyLysTyrGluThrLysLys 316
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 508 AGCAGCCGTCG---AAGTGGGACCGCTGCTCAATACCGGGATTCGATCTTTAA 564
Oy 317 AsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTrrLysAspArg 336
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 565 AATTCATTAACGACGTTTGGTGGCGCAAGGCTACTTCGACAGCGCTTCACAAACAAAGC 624
Oy 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTrrAspThrGly 356
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 625 CACCTAGCGCTTCCGCTGGAGCAGCAGCTGTCGAGCATTCGAGTATTCGATTCGACAGCGGC 684
Oy 357 ThrGlnTyrArgPheAspGlnValaIlePhePheThrIleAspProLysThrAsnGlnLeu 376
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 685 GACCGCTACCGGTTTGGCGATGTGACTTC----- 714
Oy 377 ThrThrAspProAspLysLeuProValLysArgGluLeuGlnLeuLeuThrVal 396
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 715 -----GAGAGATCGCAAAATCCGTGATGAGTATTCGCAAAACCTGGGCGCTTT 762
Oy 397 AsnMetGlyGlnAlaTrrAsnLeuGlnAlaValaIleValaIleArgAlaLeuSerAsnAspLeuIleAla 416
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 763 AAAAAGGCGCACTACTACGATCCAGCGACGACGTGGCGCAAGCTGCTTTGTCCGCG 822
Oy 417 ThrArgTrrPheAsnMetValaIleThrGluIleValaPheProGluArgGlnIleGln 436
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 823 ACGGCGTGTTAATCTCCG---GTGGTGGCGCGGAA----- 858
Oy 437 AsnAspGlnValSerPheGlnGlnSerSerSerSerArgThrGlnProAlaGlnValaAsp 456
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 859 -----TTTGATTAATCC----- 870
Oy 457 GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp 476
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 870 ----- 870
Oy 477 IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla 496
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 870 ----- 870
Oy 497 AlaLysAlaArgHisLeuTrrAspMetProAspAspArgValLeuAlaIleAsnHisAsp 516
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 870 ----- 870
Oy 517 AspGlyValaAsnArgSerIleLeuGlyArgIleSerAspAlaValaSerAlaValaArg 536
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 870 ----- 870
Oy 537 AlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeu 556
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 870 ----- 870

```

```

Oy 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysProLeuTyrValPhe 576
Db 871 -----CGTAAACG-----AAAGTGTACCGCTGCATGCGCG 903
Oy 577 ValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPLeuSerAspThrGly 596
Db 904 GTCTCTCCGCGCACCGAAACACCATGACCGCTGCGCTACCTCCACGACGCGCG 963
Oy 597 ThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGly 616
Db 964 CCGCCGCTGAAGCCTCGGGAAGAAACCGTGATGTAACCTCTACGCGCACGCTGACC 1023
Oy 617 AlaGluLeuArgLysSerGluAspLysGlyValLysLeu---TyrAlaThrLysPro 635
Db 1024 ACCACGCTGAGCTGTCTCGCGCTGACACGACGCTGACTTCAGCTATTAATGCCGCTG 1083
Oy 636 LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPhe 655
Db 1084 CTGAAGAAATCCGCTTGACGAATACTCTGTCAGGCGCGCTTAAAGCGTACC----- 1137
Oy 656 GlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluLeuSerArg 675
Db 1138 -----GATTGACAGATACCGACGACGACTCGACGACCTT-----GCGGTATCAGC 1185
Oy 676 SerIleIleGlnAsnGlyTyrPAsnArgThrTyrSerLeuArgTyrArgLeuAspLys 695
Db 1186 TTCGTGGATCTCTCCAGCGGCTGCGCGCGCATTAACCTCGCGGAGCCTCGACAC 1245
Oy 696 LeuLysThrGlnAlaProGluThrTyrPglAspLeuProValAspPheValAsnGly 715
Db 1246 TTT---ACCCAGGCA-----AACGTC 1263
Oy 716 LysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsn 735
Db 1264 ACTAACACACCATGCTGCTTATCCGGCGTGATGATACGCGTACCGCGCGCGGT 1323
Oy 736 LeuValAsnProMetArgLysTyrArgGlnArgTyrSerLeuGlnGluValGlySerSerGly 755
Db 1324 GCGCTGATCGCAGCCTGGGCGCGCTCTACGCGTACTCATGATATACCAACATGCGC 1383
Oy 756 LeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe 775
Db 1384 TGGGGCTCGACGTACTCTCCGCTTTCAGGCGCAAGCTGATCCGACGCTG 1443
Oy 776 GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGly 795
Db 1444 TACGACAAA-----CACCGCTTGTGATGCGCGGTAACTCTCGC 1482
Oy 796 TyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGly 815
Db 1483 TCGATCGAAGAACCGGATTTTCGACAGCGTCGCGCGGATCTGGCTTTCGCGGCGG 1542
Oy 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyr 835
Db 1543 GACCCAGCATTCGCGGGATATAGTACAATCATCTCACTGAAACACCAAGCCAG 1602
Oy 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnThrGluPheMetLys 855
Db 1603 TTGACCGGCGGCTCAAACTGGCGACCGGCTCGCTGAGTACCACTACAGTACCGGCG 1662
Oy 856 AspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsnGlyPheThrAsn 875
Db 1663 AAGTGGTGGGGCGGATGTTGTTGACGGCGGTGAAGCGGTGAACATATCCGCGCAG 1722
Oy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrPalaSerProValGlyGlnValArg 895
Db 1723 GATTTCACAAACCGCGCGCGGTAGCGCTGCGTCACTACCGCGTGGCGCCACAG 1782
Oy 896 ValAsp-----ValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisPhe 913
Db 1783 CTCGATTTCGCGCTTCGGTGGCGGACAAAGCAACAGCGGT-----TTACAGTTT 1833

```

```

Oy 914 PheIleGly 916
Db 1834 TACATCGT 1842

RESULT 4
US-09-913-101-1
; Sequence 1, Application US/09913101
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louise
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: Haemophilus Influenzae RD Outer Membrane
; FILE REFERENCE: BMA5368
; CURRENT APPLICATION NUMBER: US/09/913,101
; PRIORITY FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 9902880.5
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: PCT/EP00/00887
; PRIORITY FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Haemophilus influenzae

US-09-913-101-1

Alignment Scores:
Pred. No.: 8, 87e-26 Length: 1737
Score: 436.50 Matches: 142
Percent Similarity: 35.64% Conservative: 105
Best Local Similarity: 20.49% Mismatches: 235
Query Match: 9.23% Indels: 211
DB: Gaps: 14

US-09-914-168-2 (1-919) x US-09-913-101-1 (1-1737)
Oy 243 ArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer---IleGlyGlu 261
Db 217 CGGTGTTTGGTTATATATGATGATCTTCGCGCTTTGACGAAACAGCGTCAAGCGAA 276
Oy 262 ValAspValIleIleHisAspLeu-----GlyGluProValTyrIleAspTyrArgAla 279
Db 277 CGCGATTATGATGATCTCATGCTTACACCGACGCCAACAATAATTCGCGGAGCTGAT 316
Oy 280 ValGlnValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluVal 299
Db 337 GTCCAAATTGAGGGGAGACCCGCAAGATGAATAATTATATGCAATTCGCTAAACCTTA 396
Oy 300 ProLeuLeuIleGlyAspValPheHisGlyLysGlyTyrGluThrLysLysAsnLeuIle 319
Db 397 CCG---AAGATGCGCTTTTGGTTGAGCAACCAACTAGCATGATTAACAACAGCGATT 453
Oy 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPLeuAspArgSerValAsp 339
Db 454 TCACGTTTTCGATTAATGCTGGGTATTTGATGGGAGAACTTAAATTCACGCTTAGAA 513
Oy 340 ValIleLeuProAspAsnThrAlaAspValSerLeuIleGlyThrPheGlyThrGlnTyr 359
Db 514 ATGAGCCCTGAACCCATCAAGCATGTGGCGGATGTATTGATAGTGCGCTTAT 573
Oy 360 ArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAsp 379
Db 574 CATATGCGCATATTAATCTTT----- 594
Oy 380 ProAspLysLeuProValLysArgGluLeuLeuGlnGlnLeuLeuThrValAsnMetGly 399
Db 595 ---AGCCATTCAACAATTCGGGATGATTAATTAATATCTTAACAATCAATCTGG 651
Oy 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAspAspLeuIleAlaThrArgTyr 419
Db 652 GATCGTATTTAATGAATTAATTTGCGGATTTTAACGACGATTTTCCATCTCAAAATTGG 711

```

Oy	420	PheAsnMetValAsnThrCysIleValIlePheProGluArgGlnIleGlnAsnAspGln	439
Db	712	TTTGGCTCACTA-----	723
Oy	440	ValSerPheCysInserSerSerSerArgTrpGluProAlaGlnValAspCysLeuThr	459
Db	723	-----	723
Oy	460	LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro	479
Db	723	-----	723
Oy	480	IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaIalaLysAla	499
Db	724	-----TTAGTTTCAGCCCTAAATGTTAAAT-----	744
Oy	500	ArgHisLeuTyraAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal	519
Db	744	-----	744
Oy	520	AsnaArgSerIleLeuGlyArgGllEseAspAlaValSerAlaValAlaArgAlaIleLeu	539
Db	744	-----	744
Oy	540	ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg	559
Db	744	-----	744
Oy	560	LysThrProAlaAspValTyrglnSerLysValProLeuTyraValPheValAlaSer	579
Db	745	-----CATAAACGCMAAACCTGTGATGCGATGAATATTATTCCTTAATCCA	786
Oy	580	AspLysProAlaGspGlyGlnIleGlyLeuGlyTrpCylSerAspPhcSlylthArgLeu	599
Db	787	CGTAAAAAAAATGGCATGGAACTCGGTGGCTTTCTACTGATGAGCGCGCTTACCGGA	846
Oy	600	ValThrLysPheGlnHisAsnLeuIleAsnArgAspSylTyrcIlnAlaGlyAlaGluLeu	619
Db	847	CAATATAGCGCTGCACAAAACCTGGATTAATATAGCCCGTAGCATAGTTTGCGTTCAAAATCT	906
Oy	620	ArgLeuSerGluAspLysLysGlyValLys---LeuTyraIalathrlYserProLeuSerHis	638
Db	907	TATCTCTCGCACCAAAAACAACCTGTAGAGCAACTATTCGAATGCCACTGCTTAAATAAT	966
Oy	639	ProLeuAsnAspGlnLeuArgAlaIthrLeuGlyTyrcIlnGlnGluValPheGlyHisSer	658
Db	967	CCATTAAATTAATTACTATGATTTTGGCGTGGTGGGA-----GGCGAAAA	1014
Oy	659	ThrasnglyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIleIle	678
Db	1015	GAGCAT-----GATACCAGATACGAGAGTGGCTTACGCTGTGCACCGTTACGTTAATGGAT	1066
Oy	679	GlnAsnGlyGlyTrpAsnArgThrTyrserLeuArgTyraArgLeuAspLysLeuSylThr	698
Db	1069	AATGGCGATCGTTGGCAATATTTTGGCGGCACTTCGTATCGATGCAGACAGATTTT--ACA	1122
Oy	699	GlnAlaAPROProGluThrTrpClnAspLeuProValasphetaLsnnglylsProser	718
Db	1126	CAAGCGGAT-----	1133
Oy	719	GlnGluAlaLeuLeuAlaGlyValAlaIvalHisLysThrValAlaAspAsnLeuValasn	738
Db	1135	-----ATCAGTGAATAAACCTTA-----CTTCTTTAT	1166
Oy	739	PrometArgGlyTyraArg-----	744
Db	1162	CCAACTGTTGGATTACTCGTACTCGATTAACGTGTGGTCTTTCCTTCCACTTGGGCGCAT	1221
Oy	745	---GlnatrgTyrserLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla	763
Db	1222	GTGCAAAAAAATTACTTTTGATTTAAGCAAAGAAATTTGGCTATCAGAAATCTCTTTATA	1287
Oy	764	IleAlaArgAlaGlyIleSerCylValTyrserPheCysAspAsnAlaTyrglySerAsn	783

```

Db 1282 AAGTGCAGACGACTCTAGCGCGGTGGTTCGTACTATTGCAGAAAT----- 1326
Qy 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTYTILETPSerAspAsnPheAsn 803
      |||||::: ::|||::: ::|||:::
Db 1327 -----CATGCTGCGTTCGCTCGCTGAATCGGATATTTCATACAAAGGTATTGAA 1380
Qy 804 HisValProTyrIraIeuaatGpHePheAlaGlyGlyAspGlnSerILEarGlyTYTValA 823
      ::|||::: |||||::: |||||:::
Db 1381 AAAATTCGCGCTCACACGCGCTTCTTGGTGGGGAATCGTAGTGTCGGCGGTTACGCC 1440
Qy 824 HisAspSerIeuserProIleSerAspLysGlyTYTLeuThrGlyGlyValIleuAla 843
      ::: ::: ::: |||||::: |||||::: |||||:::
Db 1441 TATATAAAATTTGGCGCTTAATAATACAAATGCAAAATTTGGTGGGTGCTTCGCGATGCTT 1500
Qy 844 ValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValaPheGly 863
      ::: ::: |||||::: |||||::: |||||:::
Db 1501 ACCCACTTCTTAGATATTCATATTCAAAGTCATGCCAAATTTGGTGGGCGCAACTTTTCCA 1560
Qy 864 AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspPthLysILEgIyAlaGlyAl 883
      ||| ||| ||| ||| ::: ||| ::: ||| |||||
Db 1561 GATAGTGGATTAGCTGGCCGATTAATTACACAGCAAAAGACCTCGTTATGCCACAGCGCTT 1620
Qy 884 GlyValaArgTrrPalaserProValGlyGlnValaIargValaAspValAlaIthrcIyAllys 903
      |||||::: |||||::: |||||::: |||||:::
Db 1621 GGTGTGGCTTGGGCAATGCCACAGTTGGTGCATTAAATTTGATATATTCGCCACACCATTGCT 1680
Qy 904 GluGluGlyAsnProIleLysIleuHisPhePheILEgIy 916
      ::: ::: ::: |||||::: |||||::: |||||:::
Db 1681 GATAAAGATTAACGCAAAATTAATTCATATTTTACATCGGA 1719

```

```

US-09-913-101-3
? Sequence 3, Application US/09913101
? GENERAL INFORMATION:
? APPLICANT: Ruelle, Jean-Louis
? APPLICANT: Thomnad, Joelle
? TITLE OF INVENTION: Haemophilus Influenzae RD Outer Membrane
? TITLE OF INVENTION: Sequences Used as Vaccine
? FILE REFERENCE: Bm45368
? CURRENT APPLICATION NUMBER: US/09/913.101
? CURRENT FILING DATE: 2002-12-18
? PRIOR APPLICATION NUMBER: 9902880.5
? PRIOR FILING DATE: 1999-02-09
? PRIOR APPLICATION NUMBER: PCT/EP00/00867
? PRIOR FILING DATE: 2000-02-04
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 1731
? TYPE: DNA
? ORGANISM: Haemophilus influenzae
US-09-913-101-3

```

Alignment Scores:	
Pred. No.:	4,81e-25
Score:	427.50
Percent Similarity:	35.64%
Best Local Similarity:	20.35%
Query Match:	9.04%
DB:	6
	Gaps: 11
	Length: 1731
	Matches: 141
	Conservative: 106
	Mismatches: 235
	Indels: 21
	Gaps: 14

US-09-g14-l68-2 (1-919) x US-09-g13-l01-3 (1-1731)

Oy 243 ArgAlaValAGLYTYrTyraSPllleaspleuseSerlleIeargaSner--IlleGlyGu 261  
||| |||||||||::: ||| |||:  
Db 211 CGTGCTTGGATTATTAATCACTTCGCGCCTTTGAACAAACACGTCAAGCCAAA 270  
Oy 262 ValAsPValIIlelleHisaspLeu-----GlyGUprrovalTYrIIeasPYrArgAla 275  
|||:::|||||::: ||| ||||| |||  
Db 271 CCGGATTTATTGATTCCTCATCTTTACACACAGCGCACAAAAAATTCGCGGACTGAT 330  
Oy 280 ValGUlValArgGlyGUclUGlYlaLaaspsplysAlapheTrHThValAlaIsplVal 295





```

Oy 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGln 793
    |||||
Db 419 ACCTTCGGCCGCGC-----GGCCATCGGCTGCTGCGCGCCGCGCAG 457
Oy 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
    |||||
Db 458 GTAGCGGGAATGCGCAGCAGCACTACAGTCGATCCGCCCTCGCTGCGCTTCCTGCT 517
Oy 814 GtGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833
    |||||
Db 518 GCGCGGCGACCAAGCGCTGCTGCTACGACTACCGGACCGCTGCTGCGCGAGATTCGCA 577
Oy 834 GtYrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPhe 853
    |||||
Db 578 GCGCACAAGATCGCGCGCGCTACATGATCGCGCGAGCGTGAATCAATATCCGCTG 637
Oy 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspGlyPhe 873
    |||||
Db 638 GCGGAGCGCTGCGCTGCGCAGCTTCGTCGACGAGGCGACGCCCTCAACTCGCTGAGC 697
Oy 874 ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGln 893
    |||||
Db 698 TTCGCCGTGCATCAAGACCGGGGCTCGGCTCGCGCTGCTGCGCGGCTCGCGCGCG 757
Oy 894 ValArgValAspValAlaThrGlyValGlyGluGluGlyAsnProIleLysLeuHisPhe 913
    |||||
Db 758 TTGCGCGCTGACCTGGCCCATGCGCTCGACGACGAGCGGGT---TTCCGCGTGCACCTTC 814
Oy 914 PheIleGly 916
    |||||
Db 815 TCACATGGGG 823

RESULT 8
US-10-419-128-14966/c
; Sequence 14966, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14966
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-14966

Alignment Scores:
Pred. No.: 3,566-20 Length: 726
Score: 362.00 Matches: 82
Percent Similarity: 51.67% Conservative: 42
Best Local Similarity: 34.17% Mismatches: 90
Query Match: 7.66% Indels: 26
DB: 8 Gaps: 4

US-09-914-168-2 (1-919) x US-10-419-128-14966 (1-726)
Oy 679 GlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuTyrThr 698
    |||||
Db 672 CAGCCCGATGCTGCTGCGAGCGGTGTTCGTAACCTGATCGCGCAGAGTACAGCTG 613
Oy 699 GlnAlaProGluThrTyrPasnLeuProValAspPheValAsnGlyLysProSer 718
    |||||
Db 612 -----GGCCAGCAGCTCC 601

```

```

Oy 719 -----GlnGluAlaLeuLeuAlaGlyValAlaValAlaHisLysThrValAlaAspAsnLeu 736
    |||||
Db 600 GGGCTAGCAGAGCTTCTGTATGCGCGGATCGGCTACTCCGCTGTGGAGACGACACAG 541
Oy 737 ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeu 756
    |||||
Db 540 GTGCAGCCCGACGATGCTGCTACCGCGCTGACGATTCACAGTCAAGGGGCGAAGAGCGCTG 481
Oy 757 ValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerClyAlaLysPheGly 776
    |||||
Db 480 CTGGCCGCGACGAGCTCTCCATGTCGACGAGCGGACCGGACGCTGACACCTTCGCC 421
Oy 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyr 796
    |||||
Db 420 GGC-----GGCCATCGGCTGCTGCGCGCGCTGACGATGAGCGGA 382
Oy 797 IleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyAsp 816
    |||||
Db 381 ATCGCCACCAAGACACTACAGTCCGCGCTGCGCTGCGCTTCCTGCGCTGCGCGCGAC 322
Oy 817 GlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeu 836
    |||||
Db 321 CAGAGCGTGGTGGCTACAGACTACCGAGCGCTGTCCGAGACATTCGATGGCGACAAG 262
Oy 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAsp 856
    |||||
Db 261 ATCGCGCGCGCTGCTACATGTCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTG 202
Oy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAsp 876
    |||||
Db 201 TGGCGCGCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 142
Oy 877 ThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArgVal 896
    |||||
Db 141 ATCAAGACCGGCTGCGCTGCTGCGCGCTGCTGCTGCTGCTGCGCGCGCGCGCGCTC 82
Oy 897 AspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
    |||||
Db 81 GACCTGCGCCATGCGCTGCGCGCGCGCGCGCGGT---TTCCGCTGACCTTCCTCATGGGG 25

RESULT 9
US-10-366-683-14966/c
; Sequence 14966, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nollinger, Jork
; APPLICANT: Delouhery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PAT03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14966
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-14966

Alignment Scores:
Pred. No.: 3,566-20 Length: 726
Score: 362.00 Matches: 82
Percent Similarity: 51.67% Conservative: 42
Best Local Similarity: 34.17% Mismatches: 90
Query Match: 7.66% Indels: 26
DB: 9 Gaps: 4

US-09-914-168-2 (1-919) x US-10-366-683-14966 (1-726)
Oy 679 GlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuTyrThr 698

```

```

      611      612      613      614      615      616      617      618      619      620      621      622      623      624      625      626      627      628      629      630      631      632      633      634      635      636      637      638      639      640      641      642      643      644      645      646      647      648      649      650      651      652      653      654      655      656      657      658      659      660      661      662      663      664      665      666      667      668      669      670      671      672      673      674      675      676      677      678      679      680      681      682      683      684      685      686      687      688      689      690      691      692      693      694      695      696      697      698      699      700      701      702      703      704      705      706      707      708      709      710      711      712      713      714      715      716      717      718      719      720      721      722      723      724      725      726      727      728      729      730      731      732      733      734      735      736      737      738      739      740      741      742      743      744      745      746      747      748      749      750      751      752      753      754      755      756      757      758      759      760      761      762      763      764      765      766      767      768      769      770      771      772      773      774      775      776      777      778      779      780      781      782      783      784      785      786      787      788      789      790      791      792      793      794      795      796      797      798      799      800      801      802      803      804      805      806      807      808      809      810      811      812      813      814      815      816      817      818      819      820      821      822      823      824      825      826      827      828      829      830      831      832      833      834      835      836      837      838      839      840      841      842      843      844      845      846      847      848      849      850      851      852      853      854      855      856      857      858      859      860      861      862      863      864      865      866      867      868      869      870      871      872      873      874      875      876      877      878      879      880      881      882      883      884      885      886      887      888      889      890      891      892      893      894      895      896      897      898      899      900      901      902      903      904      905      906      907      908      909      910      911      912      913      914      915      916      917      918      919      920      921      922      923      924      925      926      927      928      929      930      931      932      933      934      935      936      937      938      939      940      941      942      943      944      945      946      947      948      949      950      951      952      953      954      955      956      957      958      959      960      961      962      963      964      965      966      967      968      969      970      971      972      973      974      975      976      977      978      979      980      981      982      983      984      985      986      987      988      989      990      991      992      993      994      995      996      997      998      999      1000

US-10-144-771-21979
: Sequence 21979, Application US/10144771
: GENERAL INFORMATION:
: APPLICANT: VENTNER, J. Craig
: TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
: FILE REFERENCE: CL001321
: CURRENT APPLICATION NUMBER: US/10/144,771
: CURRENT FILING DATE: 2002-05-15
: NUMBER OF SEQ ID NOS: 47235
: SEQ ID NO 21979
: LENGTH: 1416
: TYPE: DNA
: ORGANISM: HUMAN
US-10-144-771-21979

Alignment Scores:
Pred. No.: 4,6e-19 Length: 1416
Score: 353.00 Matches: 148
Percent Similarity: 32.63% Conservative: 85
Best Local Similarity: 20.73% Mismatches: 215
Query Match: 7.47% Indels: 266
DB: 9 Gaps: 16

US-09-914-168-2 (1-919) x US-10-144-771-21979 (1-1416)
```

```

211 AsnIleLysAlaAlaLeuGluuAspIleThrClnGlnuSerAlaMetAspLeuAsnGlySer 230
212 AAGCTTCGTCGCGACGCTTTTACGATTGAAAGTATGAAGGACCGACACCGTCGCTT 90
213 IleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr----- 248
214 CGCGACGCGTCGATGATGTCATCCGCCACCGCAAGAGGTCGAAAGCGCGGTTATTACACCG 150
215 -----AspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIle 266
216 ACCATTGAATTTGATTCCTCCGCCACCGCAAGAGGCGCGAGATTTGATGTCGCCAA 210
217 HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGly 286
218 GTCACGCGACGCGTCGCGGTTGATTTGCGGCGACCGCATGTCGATTTGCGGCGCGCG 270
219 AlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuIleGlyAspVal 306
220 CGGACCGATAAAGACTATTGAAATTCGATCTCGATCTCGCGCG--GCTATTGGCAGCGTA 327
221 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHis 326
222 CTGAACCAAGCGCGATTTATGAATAATTCAAAAGTCCTTAACACGATTCGCTGCTAAA 387
223 GlyTyrPheAspGlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThr 346
224 GGTATTTCGATTCAGATTTATCCAAAGCGCGATTCGCGCTGCGCTGCTGCTGCTGCTAAA 447
225 AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhe 366
226 GCCTTCGCGATTTGATTTATACAGTGGCGACGCTTACCGCTTTGGGACGATGACCTTT 507
227 PheThrIleAspProLysThrAsnGlnLeuThrThrAspAspLysLeuProValLys 386
228 -----GAAGATCAACAATCCGC 525
229 ArgGluLeuLeuGlnGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406
230 GATGAATAATCTGCAAAATCTGGTCCGCTTAAAGAGGCGTGAGTGAATCAAGAAAT 585
231 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlu 426
232 CTGCAGACACTGAACCGCGCTTTCGTCACCGGCTGCTTAATAATG-----#----- 633
233 IleValIleProGluArgGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSer 446
234 ----- 633
235 SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlnThr 466
236 -----CCGCTGCTGAAG--- 645
237 ValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsn 486
238 -----AATCCACTGGAA-----CAATATTAT 666
239 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
240 TTGCTGCAG----- 675
241 AspAspArgValLeuAlaIleAsnHisAspArgLysValAsnArgSerIleLeuGlyArg 526
242 -----GCGGTTTAAAGCGACT----- 693
243 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
244 -----GACCTGAACGATACCGAA 711
245 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
246 ----- 711
247 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
```



```

Db 711 ----- 711
OY 587 IIEGLYLEUGLYTPRGLYSERASPRHRLGTYHARGLEUVALDHLRYSRHEGLIUNHISA 606
Db 712 ----- 744
OY 607 LEUILEASARGSPGLYTRGILNAGLVALGLULEUARGLEUSERGLUASPLYSLS 626
Db 744 ----- 744
OY 627 GLYVALYLSLEUTYALATHRLYSPROLEUSERHISPROLEUASNPGLINLEUARGALA 646
Db 744 ----- 744
OY 647 THILEUGLYTRGILNGILNUVALRHEGLIUNHISSETHIASNGLYRHEASPLEUSERTHR 666
Db 745 ----- 759
OY 667 ARGTHRLLEUGLUNHISGLUILESERARGSERILEGLIASNGLYLYTRPRASNARGTHR 686
Db 760 ----- 774
OY 687 TYRSERLLEUGLYTRARGLEUASPLYSLSYTHRLALAPROPROGLUTHRTIRGLIN 706
Db 775 ATTAACSTGCGCTGGAGTCTGCAC----- 798
OY 707 ASPLLEUPROVALASPRHEVALASNGLYLVSPROSERGLINLUALALEU-----Ala 724
Db 799 ----- 846
OY 725 GLYVALALAVALHISLYSTHRYVALALASASPLEUVALASNPROMETARGLYTYRARG 744
Db 847 GGGGATGATATACCGCGACGGCTTCTCGTGGGCGCTGATGCACACCGGGCGACCTGC 906
OY 745 GILNAGTYRSERLLEUGLYVALGLYSERSEGLYLEUVALSERASPLASMETALALLE 764
Db 907 CAAACSTACTCTATGACACTCTCCACACAGCGCTGGGCTTCAGATTCGATTCCTCCGT 966
OY 765 ALAARGVALGLYILESERGLYALTYRSERLPHEGLYASPSALALATYRGLYSERASNARG 784
Db 967 TTCACAGCGCCACACAGCTCGATCCGACACTGATCCG----- 1008
OY 785 ALAHISGLIMETHTHRLGLYLTIEGLNALAGLYTYRILETRSERASPSANPHEASNHIS 804
Db 1009 ---CATCGTTTGTATACAGCGCGACGCTGGCTGGATTGAACCGGTGATTTCCACAA 1065
OY 805 VALPROTYRARGLEUARGRHERHEALAGLYGLYASPSGLINSERTLEARGLYTYRALAHS 824
Db 1066 GTACCGCGCGATCTGCGTTCTTCGCGGGGCGACCGCAGATTCGTGCTCAAAATAC 1125
OY 825 ASPSERLSESRPROLSESRASPLYSGLYTYRLEUTHRLGLYGLYINLALALEUVAL 844
Db 1126 AAATCTATCGCTCCGAATATACGCCAACAGGTGACTGAAGAGGCGCTGAGATTGATACC 1185
OY 845 GLYTHRALAGLYTYRASNTYRGLUPHEMELYSASPLEUARGLEUVALALPHEGLYASR 864
Db 1186 GGATGCGTGCATACACAGTACAAAGTACGCGGAAATGCTGGCGCGCGCTGTTTCGAT 1245
OY 865 IIEGLYASNALATYRASPLYSGLYRHERHNASPETHRLYSLEGLYVALGLYGLY 884
Db 1246 AGTGGCGAAGCGGTAAAGCATATTCGCGCGACGACTTAAACCGGTACCGGGCGGCG 1305
OY 885 VALARGTRALASERPROVALGLYINLARGVALASPRVALALATHRLGLYAL----- 902
Db 1306 GTGCTGGGGAATCGCGGTGGGCCAATCAATCGATTTTGCCTTACCGCGTGGGAT 1365
OY 903 LYSGLIUGLYLASNPROLLEUSLEUHNISRHEPHEILEGLY 916
Db 1366 AAAGACGAACACGGG-----TTACAGTTTTCATCGCT 1398

```

RESULT 11  
US-60-453-134-81

```

; Sequence 81, Application US/60453134
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
; FILE REFERENCE: 28335/38815
; CURRENT APPLICATION NUMBER: US/60/453,134
; CURRENT FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 588
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 81
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: H. Influenzae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a, c, g, or t
US-60-453-134-81

```

```

Alignment Scores:
Pred. No.: 6,39e-19 Length: 1365
Score: 351.00 Matches: 98
Percent Similarity: 43.87% Conservative: 63
Best Local Similarity: 26.70% Mismatches: 148
Query Match: 7.43% Indels: 58
Gaps: 8
DB: 10

US-09-914-168-2 (1-919) x US-60-453-134-81 (1-1365)
OY 566 TYRGLINSERLYSLSVALPROLEUTYRVALPHEVALALSERASPLYSPROARGSPGLY 585
Db 79 CATAAAGTAAACCGGTATAGTAATCATCTTTATCCACGTAATAAATCGCATG 138
OY 586 GILIEGLYLEUGLYTPRGLYSERASPRHRLGTYHARGLEUVALDHLRYSRHEGLIUNHIS 605
Db 139 GAACTCGGTGGGCTTCTACTGATGGCGGCGTTCCACGACAAATAGCTGGACAA 198
OY 606 ASNLEUILEASARGSPGLYTRGILNAGLVALGLULEUARGLEUSERGLUASPLYS 625
Db 199 CCGTGATTAATAGCGCGTGACATAGTTGCGTTCAATCTTATCTCTCGACCAAA 258
OY 626 LYSGLYVALYS---LEUTYRALATHRLYSPROLEUSERHISPROLEUASNPGLINLEU 644
Db 259 CAAACTCTAGAGCAACTTATCGAATCCACTCTTAAAAATCCATTAATATTACTAT 318
OY 645 ARGALATHRLLEUGLYTYRGLINGILNUVALRHEGLIUNHISSETHIASNPGLINLEU 664
Db 319 GATTTTCCGTCGCTGGGAA-----GGGGAAGAACAGAAC-----GATACC 360
OY 665 SEETHARGTHRLLEUGLUNHISGLUILESERARGSERILEGLIASNGLYTYRPA 684
Db 361 AATACGACAGTGTTCACGTTTCACGCTTACGTTATGCAATATGCGCATGTTGGCAA 420
OY 685 ARGTHRTYRSERLLEUARGTYRARGLEUASPLYSLSYTHRLALAPROPROGLUTHR 704
Db 421 TATTTTGGCGACTTCGTACGCGATACGACAGTTT---ACACAACCGGAT----- 468
OY 705 TRPGLINASPLEUPROVALASPRHEVALASNGLYLVSPROSERGLINLUALALEUVAL 724
Db 468 ----- 468
OY 725 GLYVALALAVALHISLYSTHRYVALALASASPLEUVALASNPROMETARGLYTYRARG 744
Db 469 -----ATCAGTGAATAAACCTTA-----CTTCTTATTCCAACTGTGGATTACT 513
OY 745 ----- 749
Db 514 CGTACTCGATTACGCTGCTTCCTTTCACACTTGGGCGAGTGCAGAAAATTTACTTTT 573
OY 750 GLUVALGLYSESRGLYLEUVALSERASPLASMETALALLEALARGALAGLYILE 769
Db 574 GATTTAAGCAAAAGATTGGGTATCGAATCTTCTTTATATAAAGTCGACAGCATCTAC 633

```



```
; PRIOR APPLICATION NUMBER: GB-9928197.4
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: GB-0005698.6
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: SeqWing9, version 1.02
; SEQ ID NO: 6
; LENGTH: 2379
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-148-534-6

Alignment Scores:
Pred. No.: 1,29e-11 Length: 2379
Score: 265.50 Matches: 195
Percent Similarity: 34.28% Conservative: 120
Best Local Similarity: 21.22% Mismatches: 348
Query Match: 5.62% Indels: 257
DB: Gaps: 39

US-09-914-168-2 (1-919) x US-10-148-534-6 (1-2379)
QY 116 ThrProLeuSerLeuGluGluPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
   :::::::::::::: ::::: ::::: :::::
DB 46 TCGCCTTTGGCATTTGCCACTTCACCATTCAGACATCCGTCTCGACGCTTGACGCGT 105
   :::::::::::::: ::::: ::::: :::::
QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
   ::::: ::::: ::::: :::::
DB 106 ACCGACCGCAGCACCTTATTCACCTACCTGCCGTCAAAGTCCGCGACACCTTCAAC--- 162
   ::::: ::::: ::::: :::::
QY 150 GluValValAlaProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
   ::::: ::::: ::::: :::::
DB 163 -----GACACACACGCGCAGTCCCATCATCAAAACCTGTAC 198
   ::::: ::::: ::::: :::::
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
   ::::: ::::: ::::: :::::
DB 199 GCCACCGCTTTCTTTGACGACGATCGAAGTCGAAACTGCGGACGCGGACGCTGTGAC 258
   ::::: ::::: ::::: :::::
QY 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
   ::::: ::::: ::::: :::::
DB 259 GTTATTCGAACGCCGCCACATCGCTCGCTCAACATCACCGCGCCCAAAATGCTGCAAAAC 318
   ::::: ::::: ::::: :::::
QY 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
   ::::: ::::: ::::: :::::
DB 319 GACGCG-----ATCAAGAAATAACCTCGATCGTGGCGCTGCGCGCATGCGCAATAC 369
   ::::: ::::: ::::: :::::
QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
   ::::: ::::: ::::: :::::
DB 370 TTTAATCAGCGCAGACTCAACGACGATCGCGCGCTGAAGACGAATACCTCGGCGCT 429
   ::::: ::::: ::::: :::::
QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
   ::::: ::::: ::::: :::::
DB 430 GGCMAAATCAATATCCAAATACAGCCCAAGTACCAAACTCGCCCGCAACCGCGCTGAC 489
   ::::: ::::: ::::: :::::
QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
   ::::: ::::: ::::: :::::
DB 490 ATCGACATCATCATTCGACAGGCGCAATATCGGCAAAATCACCGCATCGAATTTGAA--- 546
   ::::: ::::: ::::: :::::
QY 270 GluGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
   ::::: ::::: ::::: :::::
DB 547 GGCACCAACCAAGCTATTCCACCGCAACCAATGATCGG-----CAG 585
   ::::: ::::: ::::: :::::
QY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
   ::::: ::::: ::::: :::::
DB 586 ATGTGCTGACCGAAGCGCGCATTTGGACATGCTGACACGAAAGCAGCGGTTGACACCG 645
   ::::: ::::: ::::: :::::
QY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
   ::::: ::::: ::::: :::::
DB 646 CAGAAATTCGCCAGACATGGAAGAAAGTAACCGACTTCTACAGACAAACGCGCTACTTC 705
   ::::: ::::: ::::: :::::
QY 330 AspGlyArgTyrLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
   ::::: ::::: ::::: :::::
DB 706 GATTTCCGTAATCCTCGATACCGACATCCAAACCAACGAAGACAAACACGAGCACCATC 765
```

```
QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIle 369
   ::::: ::::: ::::: :::::
DB 766 AAATATCACCTCCACGAAGCGGACGCTTCCGCTGGGCGAAAGT-----TCGATT 816
   ::::: ::::: ::::: :::::
QY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
   ::::: ::::: ::::: :::::
DB 817 GAAGCGCAGCACCAACGAA-----GTCCCAAGGCCGGA 849
   ::::: ::::: ::::: :::::
QY 390 LeuGlnGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnValAlaArgAla 409
   ::::: ::::: ::::: :::::
DB 850 CTGGAAATCTGCTGACACGAAGCCCGCAATGTGATCCAGACCGCAGATGACCGCC 909
   ::::: ::::: ::::: :::::
QY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
   ::::: ::::: ::::: :::::
DB 910 GTTTGGGT----- 918
   ::::: ::::: ::::: :::::
QY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449
   ::::: ::::: ::::: :::::
DB 919 -----GAGATTGCAACCGCATGGCTCGGACGATACGATACAGC----- 960
   ::::: ::::: ::::: :::::
QY 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
   ::::: ::::: ::::: :::::
DB 961 -----GAAATCAGCGTACAGCGCTGCCGGAACGCCGGAACCCAA 999
   ::::: ::::: ::::: :::::
QY 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
   ::::: ::::: ::::: :::::
DB 1000 ACCGTGATTTGCTCTCTCGCATTCGAAACCGGCGGAAATTCATCGTCGAACGAATTCAC 1059
   ::::: ::::: ::::: :::::
QY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501
   ::::: ::::: ::::: :::::
DB 1060 ATCACCAGGCAACAAACAAACCCGCGACGAA-----GTGCTGGCGCGGAATTCGCGCA 1113
   ::::: ::::: ::::: :::::
QY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
   ::::: ::::: ::::: :::::
DB 1114 ATGGAATTCGCGCGCTTACGACACCTCCAAACGTCGAACGCTCAAGAGCGCGTGAAGCTT 1173
   ::::: ::::: ::::: :::::
QY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
   ::::: ::::: ::::: :::::
DB 1174 TTGGCTACTCTTCGACCAACGTCACAG-----TTTGAT 1203
   ::::: ::::: ::::: :::::
QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
   ::::: ::::: ::::: :::::
DB 1204 GCCGTCCCGCTTCCGCTAGCGCCGCAAA-----GTGCGATTGG----- 1242
   ::::: ::::: ::::: :::::
QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
   ::::: ::::: ::::: :::::
DB 1243 AACATGAGCTTGACCGAACGCTCCACC----- 1269
   ::::: ::::: ::::: :::::
QY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuT-----Gly 590
   ::::: ::::: ::::: :::::
DB 1270 -----GGCTCGCTCGACCTTGAGCGCGGC 1293
   ::::: ::::: ::::: :::::
QY 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg 610
   ::::: ::::: ::::: :::::
DB 1294 TGGGTTCCAGGATACCGGC-----TTGGTCATGTCCGCGCGCTATTCAGCAACACTG 1347
   ::::: ::::: ::::: :::::
QY 611 AspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
   ::::: ::::: ::::: :::::
DB 1348 TTCCGTAACGGGCAAGTCGCGCGCTGCGCGCTCGGAAGCAAA-----ACCACGCTC 1401
   ::::: ::::: ::::: :::::
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
   ::::: ::::: ::::: :::::
DB 1402 AACGGCTCGGTGCTTACCGACCGCTACTTACGCGACAGCGGGTCAACCTGCGGTAC 1461
   ::::: ::::: ::::: :::::
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
   ::::: ::::: ::::: :::::
DB 1462 -----GATATTTACGAAAGCCCTTCGACCCCGCAAAACATCGACACGCTCAAAACA 1515
   ::::: ::::: ::::: :::::
QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArg 690
   ::::: ::::: ::::: :::::
DB 1516 TATTAACCCACACCGCGC-----GGCGGCGCGCTAAGATGGG-TATCCCGGT 1562
   ::::: ::::: ::::: :::::
```

```

Oy 691 TTTATGLeu-AspLysLeuLysThrGlnAlaProProGlu-----ThrT 705
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1563 TACCGAATACGACCGGCTCAATTTCGGGCTGGCGGCGAACCCTGACCTCAACACTA 1622
Oy 705 pGlnAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerG 719
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1623 CAAACAAAGCAACCAACGCTATCCGACTTTATCAGGAATACGCAAAACGACGCGC 1682
Oy 719 nGluAlaLeuAlaGlyAlaValAlaHisLysThrVal-----Al 733
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1683 AGACGGCAGCTTCAAAAGCGCTGCTGTAACAAGACCGCTGCGGCGGCGCAACAAGAC 1742
Oy 733 aaAspAlaLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlyse 753
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1743 CGACACCGGCTCATGGCGCGACGCGGCGCTAC----- 1773
Oy 753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1774 -----CTGACCGGCGCTAAATCGCGCAATCGCCCTGCGGCGGCGCAAACTGCATA 1823
Oy 773 rSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1824 C-----TACTCGGCGACCCACACCAACCAACCTGCTTCTCC 1859
Oy 791 -----GlyIleGlnAlaGlyTyrIle 797
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1860 CTTAAGCAAAACCTTCACGCTGATGCTCGGCGCGAAGTGGCATTCGCGGCGCTAC-- 1917
Oy 797 eTTPSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyArgPgl 817
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1918 -----GGCAGAACCAAGAAATCCCTTCTTGAACACTTACGCGCGGCGCTGGC 1970
Oy 817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuTh 837
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1971 TTCCGTCGCGGCTACCAAAACGCGACGCTCGGCGC-----AAAGTATGACGA 2021
Oy 837 rGlyGlyGlnValLeuAlaValGly-----ThrAlaGluTyrAsnTyrGluPheMe 854
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2022 ATACGCGCAAAATCAGTACGCGCGCAACAAAAACCAACCTCTCGCGCGAGTGGCT 2081
Oy 854 t-----LysAspLeuArgLeuAlaValPheGlyAspIle 865
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2082 CTTCGCCATGCCGCTCGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2141
Oy 865 eGly----- 866
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2142 AGCGAGGCTGTGGAGCGGCAACCTATACGCGCGCGCAAAACGTAACAACAATCGGT 2201
Oy 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGly 883
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2202 TTACTCGCAAAACGCGCATTAATCCACCTTTACCAACGAATTGGCTATTCGCGCGGCG 2261
Oy 883 LglValArgTyrPheAspProValGlyGlnValArgValAspValAlaThrGlyVal 903
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2262 CGCGGTTACCTGCTGCTCGCTTTGGGTCCGATGCAATTCAGTACCGCTACCGCGCTGA 2321
Oy 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2322 GAAAAAACCGGAGACGAATCCAAACGCTTCCATTCAGCTCGCAGCAGGAGCTTC 2376

```

```

; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2391
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2391)
; OTHER INFORMATION: OMP85 cds
US-10-320-800-5

Alignment Scores:
Pred. No.: 4,42e-11 Length: 2391
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: Gaps: 40

US-09-914-168-2 (1-919) x US-10-320-800-5 (1-2391)
Oy 116 ThrProLeuSerLeuGluGlnPheAlaGlnGluSerThrGlnMetGlyIleAsnPro 135
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 46 TCGCTTGGCACTTGCCGCACTTCCACCATCCAAAGACATCCGCGCTGCAAGCTTGACGCT 105
Oy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGluProAsnSer 149
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 106 ACCGAGCGGAGTACCGTATTCACACTACCTGCGCTCAAAAGTGGCGACACTCAAC-- 162
Oy 150 GluValValProProThrLeuGluProGluLysProGluLeuIleLysArgLeuTyr 169
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 163 -----GACACACGCGCAGTCCCATCAACAAAGCGCTGTAC 198
Oy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 199 GCCACCGGCTTCTTTCGACAGCTACGCGTGGAAACCTCGGACGCGACCTTCGTAC 258
Oy 187 PheTyrGlnSerSerGlnSerGlyLysThrSerAlaIleGlySerSerHisGlnLysThr 206
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 259 GTTATCGAACGCGCCACCATCGGCTCGCTCAACATCAACGCGCAAAATGCTGCAAAAC 318
Oy 207 GluProTyrAlaAsnIleLysAlaIleGluLysP-----IleThrGlnLysSerAla 224
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 319 GACGCG-----ATTAAGAAAACCTCGAATCGTTCGCGCTGGCGCACTGCAATAC 369
Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 370 TTATATCAGCGACACTCAATCAGCAGTCCGCGCGTGAAGAAATACCTCGGCGC 429
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 430 GCGAACTCATATCAATCAATACCGCCCAAGTAAACCAACTCGCGCGCAACCGGCTGAC 489
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyLysValAspValIleIleHisAspLeu 269
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 490 ATCGACATCAGATTCAGCAGGCGCAATCCGCGCAAAATCAGCGCATGGAATTGGA-- 546
Oy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGlnGlyAlaAspAsp 289
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 547 GCGAACCAAGTCTATTCGACCGCGCAACTGATCGG-- 582
Oy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 583 -----CAATGTCCCTGACCGCAAGCGGCGCATTTGACATGG 618
Oy 307 -----PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 619 CTGACAGCAACCAACCAATTCAGCAGGAGAAATTTGCCCAAGATATGAAAAAGTACCC 678
Oy 321 AsnLaserAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspVal 340
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

Db 679 GACTTCTACCAAAATACGGCTACTTCGATTCCGATCTCCGATCCGACATCCAAACC 738  
 Qy 341 ILeuProAspThrAlaAspValSerLeuIleTyrAspThrGlyThrArg 360  
 Db 739 AACGAAGACAAAACAGACGACATCAAAATACCGTCCAGCAAGCGGACGTTCCGT 798  
 Qy 361 pheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrAspPro 380  
 Db 799 TGGGGCAAAAGTC-----TCCATCGAAGCGGACACCAAGAA----- 834  
 Qy 381 AspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGlu 400  
 Db 835 -----GTCCCAAAAGCCGAATCGAAAACTGCTGACATGAAAGCCCGGCAAA 882  
 Qy 401 AlaTyrAsnLeuGlnAlaValArgAlaIleSerAsnAspLeuLeuAlaThrArgTyrPhe 420  
 Db 883 TGGTACGAACGCCAGACGATGACCGCGGTTTGGT----- 918  
 Qy 421 AsnMetValAsnThrGluIleValPheProGluArgGlnIleGlnAsnAspGlnVal 440  
 Db 919 -----GAGATTCGAAGCCGATGGGC 939  
 Qy 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGlnSerThrLeu 460  
 Db 940 TCGGCGAGCTACGCAATACGC-----GAAATCAGCGTA 972  
 Qy 461 GluProValIleGlnThrValGluLeuThrAspGlyIleLeuMetLysIleSerPro--- 479  
 Db 973 CAGCGCTGCGCAACCGTCAAAACCGTGATTTCTCTGCAATCGCAACCGGCG 1032  
 Qy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492  
 Db 1033 CGGAAATCTACGTCACGAATACATCACCGGCAACAAACCCGCGAGAA--- 1089  
 Qy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp--- 508  
 Db 1090 ---GTCTCCGCGCTGAATTCGCAAAATGAAATCCGCACTTACGACATCCAAAGCTG 1146  
 Qy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIle 523  
 Db 1147 CAACGTTCCAAAGAGCGCGTTCGCGCTTTCGCTACTGCAACATGTCGAG--- 1197  
 Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlnSer 543  
 Db 1198 -----TTTGATGCTGTCCTCCGCTTCGCGCGCGGCAAAA--- 1233  
 Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563  
 Db 1234 -----GTGATTTTG-----AACATGAGTCTGACCGAAGCTTCCACC--- 1269  
 Qy 564 AspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
 Db 1269 ----- 1269  
 Qy 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601  
 Db 1270 ---GGTTCCTGGATTGTGACCGCGGGTGGGTTCAAGATACCGGG-----TTGGTCATG 1320  
 Qy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
 Db 1321 TCCGCGAGCGCTTCCCAAGCAACCTGTCGGTACGCGGCAAGTCCGCGCGCGCGCC 1380  
 Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
 Db 1381 TCCAGAGAGCAA-----ACCAGCTTAAACGGCTCGCTGTTACTGACCCGCTACTTC 1434  
 Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661  
 Db 1435 ACGGCGAGCGGGGTGACCGTGGGCTAC-----GATGTTACGGAAGAAAGCTTCGACCCG 1488  
 Qy 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681  
 Db 1489 CCGAAAGCATCGACGATCAACAATATAAACCACACCGGCA-----GGC 1536

Qy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696  
 Db 1537 GCAGGCACTCCGCTATGAGCGCTGCTTACCGAATACGACCGCGCTGATTTCCGTTGGTG 1596  
 Qy 697 LysThrGlnAlaProProGlnThrTyrGlnAspLeuPro-----ValAspPheVal 713  
 Db 1597 GCAGAACACTGACCGTCAACACCTTACCAACAAAGCGCCCAAAACATATGCGCACTTATC 1656  
 Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730  
 Db 1657 AAGAAATACGGCAAAACCGACGCGACAGCGAGCTTCAAGAGCGGCTGTCAAAAGT 1716  
 Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744  
 Db 1717 ACCGTGGCTGGGGCGGCAACAAACCGACGCGCTTATGGCGGACGCGGCTAC--- 1773  
 Qy 745 GlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764  
 Db 1774 -----CTGACGGCGGTGAACCGCGCAATC 1797  
 Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
 Db 1798 GCCCTGCTGGCAGCAACTGCAATAC-----TACTCGCC 1833  
 Qy 785 AlaHisGlnMetThrGly----- 790  
 Db 1834 ACCCACAACCAACCTGCTTCCCTCCGACGAAACCTTCACGCTGATGCTGGCGGC 1893  
 Qy 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808  
 Db 1894 GAAGTCGCGATTCGCGGCGGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944  
 Qy 809 LeuArgPhePheAlaGlyGlyAspArgGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828  
 Db 1945 GAAACTCTACGCGCGCGCGCTGGGTTGGTGCGGATACGAAAGCGGACGCTCGGT 2004  
 Qy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845  
 Db 2005 CCG-----AAAGCTATGACGAATACGGGCAAAATTCAGCTACGGGCAACAA 2055  
 Qy 846 ThrAlaGlnTyrAsnTyrGluPheMet-----LysAsp 856  
 Db 2056 AAGGCCACGCTCCGCGCGGCTCTTCCGATGCCCGCGCGCAAGACGCGCGACC 2115  
 Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870  
 Db 2116 GTCCGCTGAGCTGTTTGGCGGACGAGCGAGCTGTGGGACGCAAAACCTACGACGAC 2175  
 Qy 870 ----- 870  
 Db 2176 AACAGCAGTCCGCGACGCGCGGCGGCTTCAAAACATTTACGCGCGCGCAATACCAT 2235  
 Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSer 889  
 Db 2236 AAATCCACCTTTATCAACGAATGCTATTCGCGCGCGCGGCTGATGCTGCTCG 2295  
 Qy 890 ProValGlyGlnAlaArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907  
 Db 2296 CCTTATAGGCGCGGAAATTCACCTACCGCTGAAAGAAAAACCGGAAGACGAA 2355  
 Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 2356 ATCCAACGCTTCCCAATTCCAACTCGGACGACGCTTC 2391

RESULT 15

US-10-148-534-2

; Sequence 2, Application us/10148534

; GENERAL INFORMATION:

; APPLICANT: CHIRON SPA

; APPLICANT: STRATENS INSTITUTT FOR FOLKEHELSE

; TITLE OF INVENTION: 85KDA NEISSERIAL ANTIGEN

; FILE REFERENCE: P0235270

```

CURRENT APPLICATION NUMBER: US/10/148,534
CURRENT FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: GB-9928197.4
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: GB-0005698.6
PRIOR FILING DATE: 2000-03-09
NUMBER OF SEQ ID NOS: 106
SOFTWARE: SeqWin99, version 1.02
SEQ ID NO: 2
LENGTH: 2394
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-148-534-2

Alignment Scores:
Pred. No.: 4,43e-11 Length: 2394
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
Gaps: 40

US-09-914-168-2 (1-919) x US-10-148-534-2 (1-2394)
OY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
   ::::::::::::::::::::
Db 46 TCGCCTTGGCACTTGGCGACTTCACCATCCAGCATCCGCGTCGAGAGCGCTTGACCGGT 105
OY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGlnProAsnSer 149
   ::::::::::::::::::::
Db 106 ACCGAGCGCGATACCGATTTCACACTCGCCGTCGAAAGTCGGCGACACCTTACAC--- 162
OY 150 GluValValProProThrLeuGluProGluLysProGluLeuIleLysArgLeuTyr 169
   ::::::::::::::::::::
Db 163 -----GACACACACGCGAGTGGCATCATCAAAAGCTGTGAC 198
OY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
   ::::::::::::::::::::
Db 199 GCCACCGCTTCTTTGACGAGCGGTGCGAACTGCGGAGCGGCGAGCTCTGTCGACC 258
OY 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
   ::::::::::::::::::::
Db 259 GTTATCGAACGCCCCACCATTGGCTCGCTCAACATCAACGCGCGCAAAATCTCGCAAAC 318
OY 207 GluProTyrAlaAsnIleLysAlaIleLeuGluAsp-----IleThrGlnLysSerAla 224
   ::::::::::::::::::::
Db 319 GACGCC-----ATTAAAGAAACCTCGAATCGTTGCGGCTGCGGCAATCGCAATAC 369
OY 225 MetLysP-----LeuAsnGlySerIleProArgLeuArgGln----- 236
   ::::::::::::::::::::
Db 370 TTTAATCAGCGGACATCAATCAGCGAGTCCGCGCTGAAAGAAAGAAATACCTCGGCGCC 429
OY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
   ::::::::::::::::::::
Db 430 GCGAAACTCATATATCAAAATGACGCGCAAAAGTAAACCAAACTCGCGCGCAACCGGTGCAC 489
OY 250 IleAspLeuSerIleIleArgAsnSerIleGlyValAlaAspValIleIleHisAspLeu 269
   ::::::::::::::::::::
Db 490 ATCGACATCAGATGAGAGGAGGCAAAATCCGCAAAATCAGCGAATCGAATTGGA--- 546
OY 270 GlyLysProValTyrIleAspTyrArgAlaValGluValArgGlyGlnGlyAlaAspAsp 289
   ::::::::::::::::::::
Db 547 GCGCAACCAAGTCTATTCGCGCGCAAACTGATGCGG----- 582
OY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
   ::::::::::::::::::::
Db 583 -----CAATGTCCCTGACCGCAAGGCGGCGCATTTGGCATGG 618
OY 307 -----PheHisIleGlyLysTyrGlnGluThrLysLysAsnLeuIleGlu 320
   ::::::::::::::::::::
Db 619 CTGACACGCAACCAATTCACACGAGCAAAATTTGCCCAAGATATGCAAAATTAACCC 678
OY 321 AsnAlaSerAlaGlnHisGlyTyrThrAsnArgLysArgThrLeuAspArgSerValAspVal 340
```

```

Db 679 GACTTCTACCAAAATTAACGCGTACTTCGATTTCCGATCTCCATCCGACATCCCAACCC 738
   ::::::::::::::::::::
OY 341 IleLeuProAspAsnThrAlaAspValSerIleIleTyrAspThrGlyThrGlnTyrArg 360
   ::::::::::::::::::::
Db 739 AACGAAGACAAAACCAAGACGACCATCAAAATACCGCTCCACGACGCGGACCTTCCGT 798
OY 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
   ::::::::::::::::::::
Db 799 TGGGGCAAACTC-----TCCATCGAAGGCGACACCAACGAA----- 834
OY 381 AspLysLeuProValLysArgGluLeuGlnLeuLeuThrLysAlaAsnMetGlyGlu 400
   ::::::::::::::::::::
Db 835 -----GTCCCAAGCCGCAACTGGAAAACTGCTGACATGAAGACCCGGCGAAA 882
OY 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
   ::::::::::::::::::::
Db 883 TGGTACCAACGCCACGATGACCGCGTTTGGT----- 918
OY 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
   ::::::::::::::::::::
Db 919 -----GAGATTGCAACCCGATGCGC 939
OY 441 SerPheGluInSerSerSerArgTyrThrGluProAlaGlnValAspGluSerThrLeu 460
   ::::::::::::::::::::
Db 940 TCGGACGCTACGCAATACAC-----GAAATCAGCGTA 972
OY 461 GluProValIleGluThrValGluLeuThrAspGlyLysLeuMetAspIleSerPro--- 479
   ::::::::::::::::::::
Db 973 CAGCGGTGGCGCAACGCTGAACCAACCGCTGATTCGTCGTCGACATCGAACCGGCG 1032
OY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
   ::::::::::::::::::::
Db 1033 CGGAAATCTACGTCGAACGAATACATCAACGCGGCAACCAACCCGCGAGAA--- 1089
OY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
   ::::::::::::::::::::
Db 1090 ---GTGTCGCGCGTGAATTAACGCCAAATGGAATCCGACCTTACGACACTCCAACTG 1146
OY 509 -----ArgValLeuAlaIleAsnHisAspGlyValAlaAsnArgSerIle 523
   ::::::::::::::::::::
Db 1147 CAACGTTCCAAAGCGCGTCGACGCTTTGGGCTACTTCGACATGTGCCAG----- 1197
OY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
   ::::::::::::::::::::
Db 1198 -----TTGATGCTGTCCCGCTGCGCGGACGCGCGCAAA--- 1233
OY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
   ::::::::::::::::::::
Db 1234 -----GTGCAATTG-----AACATGAGTGTGACCGAAGCTTCCACC----- 1269
OY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
   ::::::::::::::::::::
Db 1269 ----- 1289
OY 584 AspGlyGlnIleGlyLeu-----GlyTyrPylSerAspThrGlyThrArgLeuValThr 601
   ::::::::::::::::::::
Db 1270 ---GGTCCCTGGATTTGAGCGCGGTTGGGTGAACATACCGGG-----TTGGTCA TG 1320
OY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
   ::::::::::::::::::::
Db 1321 TCGCGAGCGGTTTCCCAAGACAACTGTTCGTTAGGCGAAGTGGCGGCACTCGCGGCC 1380
OY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
   ::::::::::::::::::::
Db 1381 TCCAGGAGCAAA-----ACCACGCTTAACGCGTCCGTGTACTAGTACGCGCTACTTC 1434
OY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlnHisSerThrAsnGly 661
   ::::::::::::::::::::
Db 1435 ACGGCAACGCGGTCGAGCTGGCTAC-----GATGTTTACGAAACGCTTGCACCGG 1488
OY 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681
   ::::::::::::::::::::
```

```

Db 1489 CGCAAGCATCGACGACATCAAAATATMAAACGACGACGCA-----GGC 1536
Qy 682 GLYTPASArG-----ThrTyrSerLeuArgTyrIrgLeuAspLysLeu 696
Db 1537 GCAAGCATCCGCAATGAGCTGCTGTACGAATACGACCGCGTGAATTTGGTTGGTG 1596
Qy 697 LysThrGlnAlaProProGluThrTyrGlnAspLeuPro-----ValAspPheVal 713
Db 1597 GCAGAACACCTGACCTCAACACCTACAAAGCGCCCAAAACATATGCGCACTTATAC 1656
Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyAlaValHisLys 730
Db 1657 AAGAAATACGGCAAAACCGACGCGACGCGAGCTTCAAAAGCTGGCTGTACAAAGCT 1716
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 1717 ACCGTGGCTGGGGGGCGCAACAAACCGACGCGCTTATGGCGGCGGCTAC--- 1773
Qy 745 GlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAlaLe 764
Db 1774 -----CTGACGGGCGGTGAACGGCGCAATC 1797
Qy 765 AlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 1798 GCCCTGCTGGCGCAAACTGCAATAC-----TACTCCGCC 1833
Qy 785 AlaHisGlnMetThrGly----- 790
Db 1834 ACCCAACACCAACCTGCTTCTCCCTGAGCAAAACCTTCACGCTGATGCTCGGCGC 1893
Qy 791 -----GlyIleGlnAlaGlyTyrIleTyrPheAspAsnPheAsnHisValProTyrArg 808
Db 1894 GAAGTGGCATTCGGCGGCGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944
Qy 809 LeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrIleHisAspSerLeuSer 828
Db 1945 GAAACTTCTACGCGCGCGCTGGTGGTGGGATACGAAAGCGGACGCGCTGGT 2004
Qy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
Db 2005 CCG-----AAAGTCATGACGAATACGGGAAATAATCATGACGCGGCGCAACAA 2055
Qy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 2056 AAGCCACAGCTCCCGCGAGCTCTTCCGATGCCGCGCGCAAGACGCGCGAC 2115
Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 2116 GTCCGCTGAGCTGTTTGGCGACGACGCGTGGGCGCAAAACCTACGACGAC 2175
Qy 870 ----- 870
Db 2176 AACAGCAGTTCGCGACGCGCGCGAGGTTCAAAACATTTACGCGCGCATACCCAT 2235
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValArgTyrPalaSer 889
Db 2236 AATCCACCTTACCAACGAATTCGCTATTCGCGCGCGCGGTACTGCTCTCG 2295
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 2296 CTTTAGGCCCGGAAATTCACCTACCTACCCGCTGAAGAAACCGGAAGACGAA 2355
Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 2356 ATCCACAGCTTCCATTCCACTCGCAGACGCTC 2391

```

```

; FILE REFERENCE: P02352720
; CURRENT APPLICATION NUMBER: US/10/148,534
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: GB-9928197.4
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: GB-0005698.6
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: SeqMan99, version 1.02
; SEQ ID NO 10
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-148-534-10

Alignment Scores:
Pred. No.: 4.43e-11 Length: 2394
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
Gaps: 40
DB: 9

US-09-914-168-2 (1-919) x US-10-148-534-10 (1-2394)
Qy 116 ThrProLeuSerLeuGlnGluPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
Db 46 TCGCTTTGGCACTTCCGCGACTTCACCATTCGAAGACATCGCGCTCGAAGCTTGACGCT 105
Qy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGluProAsnSer 149
Db 106 ACCGACCGCAGTACCGTATTCACACTACCTCCCGTCAAAAGTCGCGCACCTCAAC--- 162
Qy 150 GluValValAlaProProThrLeuGlnProGluProGluProGluProGluProGluPro 169
Db 163 -----GACACACAGCGAGTCCATCATCAAAACCTGTAC 198
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
Db 199 GCCACGGGTTCCTTGGACGACGTACCGGTGCAAACTGCGGACGCGGACGCTCGTGAC 258
Qy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 259 GTTATGAAACGCCACCATCGCTGCTCAATCACCGCGCAAAATATGCTGCACAAAC 318
Qy 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
Db 319 GACGCC-----ATTAGAAACCTCGAATCGTTCGGCGTGGCGGACGTGCAATAC 369
Qy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 370 TTTAATCAGCGACATCATTCAGCGAGTCCGCGCTGAAAGAAATATCTCGGGCGC 429
Qy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
Db 430 GCCAACTCAATATCCAAATACGCGCCAAAGTCAACCAACTCCCGCAACCGCGTGCAC 489
Qy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
Db 490 ATCGACATCAGATTGACGAGGCGCAATCCGCCAAATATCACCGACATCGAATTTGAA--- 546
Qy 270 GlyLnuProValTyrIleAspTyrArgAlaValAlaValAlaArgLysGluAlaAspAsp 289
Db 547 GGCACACAGTCTATTCGACCGCAAACTGATGGG----- 582
Qy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 583 -----CAGATGCTCCTGACCGCAAGCGCGCATTTGACATGG 618
Qy 307 -----PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
Db 619 CTGACACGAAGCAACCAATTCACGAGCAAGAAATTTGCCCAAGACATGAGAAAGTAAAC 678

```

QY 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspVal 340  
 Db 679 GACTTCTACAGAACAGCGCTACTTCTGATTTCCGATCTCCATCCGACATCCAAACC 738  
 QY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrArg 360  
 Db 739 ACGAAGACAAACCAACGACGACCATCAAAATCACCTCCACGACGCGGCGCTTCGT 798  
 QY 361 PheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380  
 Db 799 TGCGGCAAAAGTC-----TCCATCAAGGCGACACCAAGAA----- 834  
 QY 381 AspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGlu 400  
 Db 835 -----GTCCCAAGCGCAACTGGAAAACTGGTGAACCATGAAGCCCGGCAAA 882  
 QY 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420  
 Db 883 TGGTACGACGCGCGAGATGACCGCGCTTTGGGT----- 918  
 QY 421 AsnMetValAsnThrGluIleValPheProGluArgGlnGlnIleGlnAsnAspGlnVal 440  
 Db 919 -----GAGATTCAAGAACCGCATGGCC 939  
 QY 441 SerPheGlnGlnSerSerSerArgTrpGluProAlaGlnValAspGluSerThrLeu 460  
 Db 940 TCGGACGCTACGATACAGC-----GAATACAGCGTA 972  
 QY 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479  
 Db 973 CACGCGCTCCCAACCGCCGAAACCAAAACCGTCGATTTCCGTGCAATCGAACCGGGCC 1032  
 QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492  
 Db 1033 CGGAATCTACGTCACAGCAAGAAATCCACATCCCGCAACCAAAACCGCGAGCAA--- 1089  
 QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508  
 Db 1090 ---GTCTGCGCCCGCGCAATTCGCAATCGCGCTTACGACACCTCCAAAGCTG 1146  
 QY 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523  
 Db 1147 CAACGCTCCAAAGACGCGCTGAGCTTTGGGTACTTCCGACACGTCACG----- 1197  
 QY 524 LeuGlyArgIleSerAspAlaValSerAlaAlaArgAlaIleLeuProAspLysSer 543  
 Db 1198 -----TTTGATGCCGCTCCCGCTTCCGCGACACACCGACAAA--- 1233  
 QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563  
 Db 1234 -----GTGATTTG-----AACATGAGCTGACCGACGTTCCACG----- 1269  
 QY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
 Db 1269 ----- 1269  
 QY 584 AspGlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr 601  
 Db 1270 ---GGCTCGCTCGACTTGACGGCGGCTGGTACAGATACCGGC-----CTGGTCAAG 1320  
 QY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621  
 Db 1321 TCCGACGCGCTTCCCAACACACACTGTCGTACGGGCAAGTCGCGCGCTCGCGCC 1380  
 QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641  
 Db 1381 TCACCAACGAAA-----ACCAGCTCAACGGCTGCTGTGTTACGACCGCTACTTC 1434  
 QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661  
 Db 1435 ACGGACGACGGGTCAGCTGGGTAC-----GATGTTTACGGAAGACCTTCACCGCC 1488  
 QY 662 PheAspLeuSerThrArgThrLeuGlnHisGlyIleSerArgSerIleIleGlnAsnGly 681

Db 1489 CGCAAGCATCGACGACGATCAAAATATAAACCACACCGCA-----GGC 1536  
 QY 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrIleLeuAspLysLeu 696  
 Db 1537 GCAGGACATCCGATGAGCGCTGTTACCGAATACAGCCGCTGATTTCCGTTGGGTG 1596  
 QY 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713  
 Db 1597 GCAGAACCTGACCGTCACACCTACCAACAAAGCGCCAAACACTATGCGCACTTATC 1656  
 QY 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730  
 Db 1657 AAGAAATACCGCAAAACCGACGACACGAGCGAGCTTCAAGCGCTGGTACAAAGT 1716  
 QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgIleTyrArg 744  
 Db 1717 ACCGTGCGCTGGGGCGCCAAACCAACCGACGCGCTTATGGCGGACCGCGCTAC--- 1773  
 QY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764  
 Db 1774 -----CTGACGGGCGCTGAACGCGCAAAATC 1797  
 QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784  
 Db 1798 GCCCTGCCCGGACGCAAACTGCAATAC-----TACTCCGC 1833  
 QY 785 AlaHisGlnMetThrGly----- 790  
 Db 1834 ACCCAACCAACCAACCTGTTCTTCCCTTAAGCAAAACCTTACACCGTATCGTGGCGGC 1893  
 QY 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808  
 Db 1894 GAACTCGCATTTGGCGGCGGTAC-----GGCAACCAAAAGAAATCCCTTTT 1944  
 QY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828  
 Db 1945 GAAACTTTACGGCGGCGGCTGGTGGTGGTGGCGGATACGAAACGGGACGCTCGGT 2004  
 QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845  
 Db 2005 CCG-----AAAGTATGACGAATACGCGCAAAATACACTACGGCGGGAACAA 2055  
 QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856  
 Db 2056 AAAGCCAACTCTCGCGGACGCTCTTCCGATGCCCGGCGGAAAGACGCGGCAC 2115  
 QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870  
 Db 2116 GTCCGCTGACGCTGTTGCCGACGACGACGCGTGGGACGCAAAACCTACGACGAC 2175  
 QY 870 ----- 870  
 Db 2176 AACAGCACTTCCGCGACCGCGGCGAGGTTCAAAACATTTACGGCGCGGCAATACCAT 2235  
 QY 871 ---LysGlyPheThrAspAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889  
 Db 2236 AAATCCACCTTTACCAACGAATTCGCGATTCGCGCGGCGGCTTACCTGGCTCTCG 2295  
 QY 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907  
 Db 2296 CTTTAGGCGCGGATGAATTCAGTACGCTACCGCGCTGAAAGAAACCGAAGACGAA 2355  
 QY 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 2356 ATCCAAACGCTTCAATTCCAACTCGGACGACGCTTC 2391  
 RESULT 17  
 US-09-689-065A-1  
 ; Sequence 1, Application US/09689065A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pfizer Products, Inc.  
 ; TITLE OF INVENTION: LAMSONIA INTRACELLULARIS PROTEINS AND RELATED METHODS AND



```

: TITLE OF INVENTION: MATERIALS
: FILE REFERENCE: 3153, 00187/PC10589A
: CURRENT APPLICATION NUMBER: US/09/689,065A
: PRIOR FILING DATE: 2000-10-12
: PRIOR APPLICATION NUMBER: US Prov. 60/160,922
: PRIOR FILING DATE: 1999-10-22
: PRIOR APPLICATION NUMBER: US Prov. 60/163,858
: NUMBER OF SEQ ID NOS: 112
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO: 1
: LENGTH: 6617
: TYPE: DNA
: ORGANISM: Lawsonia intracellularis
US-09-689-065A-1

Alignment Scores:
Pred. No.: 7,85e-09 Length: 6617
Score: 238.50 Matches: 198
Percent Similarity: 34.72% Conservative: 137
Best Local Similarity: 20.52% Mismatches: 391
Query Match: 5.05% Gaps: 239
DB: 7 Gaps: 38

US-09-914-168-2 (1-919) x US-09-689-065A-1 (1-6617)
Oy 50 AsnGlnAlaIyAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArg 69
Db 3764 AATGCTGCTTCAAAAGACATCTTCTATTGTGGTTCTCCCAATTCAAATTAATGGCTCA 3823
Oy 70 LeuAsnAlaIaGlyLeuAsnAlaLys---ProGlnSerGlnAlaLeuAspValAlaAsn 88
Db 3824 TCAATGATGAAGAGTTTCAACACAGACACCAATGCTCTTGCAACTGCATTAAAGAT 3883
Oy 89 PheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAsp 108
Db 3884 -----AAGGCA 3889
Oy 109 MetSerValIleGluGluThrThrProLeuSerLeuGluLeuPheAlaGlnIleuSer 128
Db 3890 TTTCGTGTCATCCCTAATTAATCTGCATTAACTTT-----CTATTATAACAAAATATAC 3943
Oy 129 ThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn 148
Db 3944 TCCCAACTTAATATTCTCTACTGCAAAAAGGTA-----GCTCAACAACATCCATC 3991
Oy 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168
Db 3992 GCTGACTATGTAGTA----- 4006
Oy 169 TyrAlaIaGlyLeuPheAsnAspGlyValAlaLysValProArgLeuLysAlaLysPheTyr 188
Db 4007 TACGGCAGTTTCAATCAACACAGGTGAAT-----TTTACT 4042
Oy 189 GlnSerSerGlnSerGlyIleThrSerAlaIleGlySerSer-----HisGlnLys 205
Db 4043 ATTGATAGTAGCCTTATTCATGTACAGGTGTAGCATCTGCACGCCCATTTATACATAGAA 4102
Oy 206 ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMet 225
Db 4103 AAACCAAAATTTAATGAGCTAATATTGCTGTACAGAACTGCTGCAAGCTAATAGT--- 4159
Oy 226 AspleuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaVal 245
Db 4160 -----AATGCCCTTATTAAGAAA---AACACTATTGCTGATGTACGTATTTCATGGGCTT 4210
Oy 246 GlyTyrTyrAspIleAspSerIleIleArgAsnSerIleGlyIleValAspValIle 265
Db 4211 AAAGTTCTTGATCTGATGTATTCCTTACACACATCATTAATTAAGGAGATCTACTACT 4270
Oy 266 IleHis-----AspleuGlyLysIleProValTyr 274
Db 4271 GATCATGCCCAAAATTAATGACAGAAATCAAAAATAATGCGAATTAGGA-----TATTTT 4324

```

```

Oy 275 IleAspTyrArgAla---ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThr 293
Db 4325 ACTGATGCTCTGTCAAGATTAAGAAAACGGGGAAGA-----CGATTACTTGTGA 4375
Oy 294 ThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGlyTyrGlu 313
Db 4376 TTTCAGTACAGAAAGAGCTAAATTAACAGATGTTGTTGTCACAGGCTCAAAAGCTGTA 4435
Oy 314 ThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyr 333
Db 4436 ACTATGATAAACATTTCTGCTGCAATGAGTCTAAAGAAAGA----- 4477
Oy 334 LeuAspArgSerVal-----AspValIleLeuProAspAsnThrAlaAspValSerLeu 351
Db 4478 -----TCAGTTATTAGTATGACATATGTCCTCCAGATATTCAAAAATTTACCGAC 4528
Oy 352 IleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspPro 371
Db 4529 CTTATAGAAAAGAGGCTACTATCTCGCTGAA-----GTTAATTAATGAATAAAGAG 4582
Oy 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu 391
Db 4583 AAAGAAAATATCTTCTGTGCA----- 4606
Oy 392 GlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSer 411
Db 4607 CTATTGTTACAGTAATATAGAGGAAAAACTTTATATTAAAGATGTCGGAATTGACAGA 4666
Oy 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431
Db 4667 CTGGAACAAATAAAGCTTAA-----ACTTTAAAAAAGATGTACATTAAACAGAA 4717
Oy 432 Arg-----GluGlnIleGlnAsn 437
Db 4718 CGTAATTTTTCATGCTTACTGACAGAGGTGATTAACGTGAAGAAATATCTTGAACT 4777
Oy 438 AspGlnValSerPheGluGlnSerSerSerArgTyrGluProAlaGlnIleAspGlu 457
Db 4778 GACTCTATAGCAATC-----TCTGCTATGCCATGAATCATGCGTATGTAGTATAT 4828
Oy 458 SerThrLeuGluPro-----ValIleGluThrValGluLeu 469
Db 4829 CAAGTTGCTTCAACCTGAATACATTCATGAAGAAAGAAATGTTATTCATTAGAGTA 4888
Oy 470 ThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln 489
Db 4889 AAAGAAAGGTAAAGCCCTAATAATAGCAAAAATAGACTTTAAAGAGACATCTTATTGAGACA 4948
Oy 490 AspLysLeuAsnLeuValAlaIaLysAlaArgHisLeuTyrAspMetProAspAspArg 509
Db 4949 AATGAACACATCTCTTAATTAACAAAATTTGATGATCAATTAACATAGCAGATTTT 5008
Oy 510 ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAsp 529
Db 5009 TCTCTTCTGTTATGCAAGATGAT-----GTAAAGCAATTAACAGATTTTATTCAGAT 5062
Oy 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
Db 5063 TATGTTATGATCTTGTCTAAGATAGATCTT---GAACAACCAAAAATGAAGAAGATGCA 5119
Oy 547 -----ValIleAspLeuProGluArgThrAlaLeu 556
Db 5120 ACAATGATGATCTTCTTCTTATGATTAACAAAACAAAAGCTTCTTCTTCTGTAAGATAT 5179
Oy 557 -----AlaAsnArgLysThrProAlaAspValTyr 566
Db 5180 GTTGAAGAAATATCTGCTAGAGATATATTATCTCGGATTAATTAAGCCTTGCTGAT 5239
Oy 566 ----- 566
Db 5240 GGAGATCTTTTAAATGTCACACATCTCGACGCTTAATGAATGCCTTAAACGCGCTTGCC 5299

```

```
OY 567 -----GlnSerLysValProLeu 573
Db 5300 TATTTAACCAAGTAGATACAGATACAGTGCCTACAGGAAAGATGATGATGATCTA 5359
OY 574 TyrValLeuValAlaSerAspLysProArgAspGlnIleGlyLeuGlyTyrGlySer 593
Db 5360 CTGTAAAGCTT-----CAAGAAAGCTCGAAGAGTGCATACAGAGTGGTGGTAC 5413
OY 594 AspThrGlyThrArgLeuValThrLys-----PheGlnHisAsnLeuIleAsnArg 610
Db 5414 TCACACATCTTAATTTGGTGTTCAGGAAGATCTCGAAGAAACTTATGGGGA 5473
OY 611 AspGlyTyrGlnAlaGlyAlaGlyLeuArgLeuSerGlyAspLysGlyValLysLeu 630
Db 5474 ---GGTATATTTAAAGTATGAGGTTTATTTCTAGTAAGTCACTCTTCTTGATCTT 5530
OY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
Db 5531 TCTTTTACCACATCCCTCGTGTATGATACAGAC----- 5563
OY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeu----- 664
Db 5564 -----TTTGGCTTTAGTATACATTTTACGCTACGAGATGATGGATGAC 5611
OY 665 -----SerThrArgThrLeuGlnHisGlyIleSerArgSerIle 677
Db 5612 TTCCGTAATAAAACTTATGAGATACATACAGTCTATTTCCACCCATATA----- 5659
OY 678 IleGlnAsnGlyTyrPheAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLys 697
Db 5660 -----GGAGAAATATTCATCATCTATCTTGTGGCTACGAATGATCAATATCTCT 5707
OY 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValAsnIleLysPro 717
Db 5708 CTATATGATATTCATCTACACAGCAGCAGCTCTTATCTTGACATAT---CAAGGAAAT 5764
OY 718 SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 5765 ATTTCTAGTCTAGTAGTGTGTCTTACTTTGATTTCTACA-----GACAGTCTGAG 5818
OY 738 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuVal 757
Db 5819 AGACCACTTAAGGGCATTTGCAAAACATAATTTGATATGAGAGTGTGTCTGTGT 5878
OY 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 5879 GGTAAATGATTAAGCTTCTCAAGCCATTTGCTGAACCTACAAGATTTTACTCAAT 5932
OY 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGlyTyrIle 797
Db 5933 -----TCAGAAAGCTAAAAACCATATATATGATTCGGCTGACAGTGCAGCT 5983
OY 798 TrpSerAspAsnProPheAsnHisValProTyrArgLeuArgPhePheAlaGlyAspGln 817
Db 5984 TATAAGATATGTAATAAAACCTGCGCAGTATTTGACCGATTTTATTTGATGATAGAT 6043
OY 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys---GlyTyrLeu 836
Db 6044 AGTATAGAGATATGATACAGAAAGATCTTGACACAAAGATCCGCTTGGAGATGAA 6103
OY 837 ThrGlyGlyGlnValLeuAlaValAlaGlyThrAlaGlyTyrAsnTyrGlnPheMetLysAsp 856
Db 6104 ATTTGGTGTATAGAGATGGCTTTTCTTAACCTGAGATATATTTGGACATTTCCACCCAGAG 6163
OY 857 LeuArgLeuAlaVal-----PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThr 874
Db 6164 CTAGCTCTTCATTAAGTTCATGATGATAGATTCGCAACAGATTCCTGTCACAAC 6223
OY 875 AsnAsp-----ThrLysIleGlyAlaGlyValGlyValAlaGlyTyrPheAspPro 890
Db 6224 TCTAACCATTTCTTAACCTCAACAAATCATATGGCCTTGAACCTTGGCTGCTTACCA 6283
OY 891 ValGlyGlnValArg 895
```

```
Db 6284 ATGGAGATTTGCCA 6298
RESULT 18
US-09-689-065B-1
; Sequence 1, Application US/09689065B
; GENERAL INFORMATION:
; APPLICANT: Pfizer Products, Inc.
; TITLE OF INVENTION: LAWSONIA INTRACELLULARIS PROTEINS AND RELATED METHODS AND
; TITLE OR INVENTION: MATERIALS
; FILE REFERENCE: 3153, 00187/PC10589A
; CURRENT APPLICATION NUMBER: US/09/689, 065B
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: US Prov. 60/160,922
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: US Prov. 60/163,858
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 6617
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-09-689-065B-1

Alignment Scores:
Pred. No.: 7,85e-09 Length: 6617
Score: 238.50 Matches: 198
Percent Similarity: 34.72% Conservative: 137
Best Local Similarity: 20.52% Mismatches: 391
Query Match: 5.05% Indels: 239
Db: 7 Caps: 38

US-09-914-168-2 (1-919) x US-09-689-065B-1 (1-6617)
OY 50 AsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnIleIleGlnAlaArg 69
Db 3764 AATGCTGCTTCAAAAGACAGATCTTCTATTTGGTGTCCCTTCAATTAATGCTCA 3823
OY 70 LeuAsnAlaIleGlyLeuAsnAlaLys---ProGlnSerGlnAlaLeuAspValAlaAsn 88
Db 3824 TCAATGATGAAGAGTTCACAAACAGAACATGCTTCTTGACACATTAAGAT 3883
OY 89 PheAspAspGlnSerProIleSerArgIleGlyGlnSerProProLeuGlyLeuAsp 108
Db 3884 -----AAGGA 3889
OY 109 MetSerValIleGlnGluThrThrProLeuSerLeuGlnGluLeuPheAlaGlnIleUser 128
Db 3890 TTTCGTGTCATCCCTAATAATCTGCATTAAATCTT-----CTATATTAACCAAAATATC 3943
OY 129 ThrGlnMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGlnIleProAsn 148
Db 3944 TCCCACTTAATATTTCTACTGCAAAAAAGTA-----GCTCAACAATCCAT 3991
OY 149 SerGlnMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGlnIleProAsn 168
Db 3992 GCTGACATGTAFTA----- 4006
OY 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188
Db 4007 TACGGCAGTTTCANATCAACAGCTGAATAAT-----TTTACT 4042
OY 189 GlnSerSerGlnSerGlyGlnThrSerAlaIleGlySerSer-----HisGlnLys 205
Db 4043 ATTTGATATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4102
OY 206 ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMet 225
Db 4103 AAACCAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4159
OY 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrValLeuValAlaAlaArgAlaVal 245
```

Dh 4160 -----AATGGCTTATAAAGAA---AACACTATTGCTGATGTACGTATTCAGGGCTT 4210  
Oy 246 GATYTYTAspIleaspleuSerIleIleArgAsnSerIleGlyValAlaSpvalIle 265  
Dh 4211 AAGATTCTTGATCTCGATGATTAATCTTACACGACTGACTTAATTAAGGAGATCTACT 4270  
Oy 266 ILeHis-----AspleuGlyluProValIlyr 274  
Dh 4271 GATCATGCCAAATTAATGCAGAAATCAAAAATATGGAAATTAGA-----TATTTT 4324  
Oy 275 ILeasPTyArgAla---ValGluValArgGlyGlyValAlaAspAspIlySerIle 293  
Dh 4325 AGTGAATGCTCTGCAGATTAAGAAAGCGGGAAGA-----CGATTACTTGA 4375  
Oy 294 ThrValAlaAspGluValProleuLeuIleGlyAspValPheHisIleGlyIlyrGlu 313  
Dh 4376 TTTACTGTCAAGAAAGCTTAAATTAACAGATGTTGTTGTCAGGCTCAAAAGCTGA 4435  
Oy 314 ThrIlyAsnIleuIleGluAsnAlaSerAlaGluHisGlyTyPheAspGlyArgTyP 333  
Dh 4436 AGTATCGATAACATCTCTGCTGCAATGAGTCTAAAAAGA----- 4477  
Oy 334 LeuAspArgSerVal-----AspValIleleuProAspAsnThrAlaSpvalSerLeu 351  
Dh 4478 -----TCAGTTATTAGTATGATGACTATGCTCCCAAGATATTCAAAAAATTACCGAC 4528  
Oy 352 ILeTyAspThrGlyThrGlnTyArgPheAspGluValAlaPhePheThrIleAspPro 371  
Dh 4529 CTCTATGAAAAGAAAGCGTACTATCTCGCTGA-----GTTAATTATGAATAAAAAGAG 4582  
Oy 372 LysThrAsnGlnleuThrThrAspProAspLysleuProValLysArgGluLeuGlu 391  
Dh 4583 AAGAAGAAATACCTCTCTGCA-----ACA 4606  
Oy 392 GlnleuLeuThrValAlaAsnMetGlyGluAlaTyArgAsnGlnAlaValArgAlaLeuSer 411  
Dh 4607 CTATTGTTTAAAGATAAAGGAAAGAACTTTATTAAGATGTCGCAATTGAAGA 4666  
Oy 412 AsnAspLeuIleAlaThrArgTyPheAsnMetValAsnThrGluIleValPheProGlu 431  
Dh 4667 CTGGAACAATGAAGCTAAA-----ACTTAAAAAAGAGTTACCATTAACAGAA 4717  
Oy 432 Arg-----GluGlnIleGlnAsn 437  
Dh 4718 CATAATTTTTTATCATGTTTACTGGAACAGGTGATTACGTGAAGATATCTTGAAGCT 4777  
Oy 438 AspGlnValSerPheGluInserSerSerArgThrGluProAlaGlnValArgGlu 457  
Dh 4778 GACTCTATAGCAATC-----TCTGCCATATGCCATGAATCATGGCTATGATATTT 4828  
Oy 458 SerThrleuGluPro-----ValIleGluThrValGluLeu 469  
Dh 4829 CAAGTGGCTTACACCTGAATACATTCATGAAGAAAGAAATTTTATTCATTTAAGTA 4888  
Oy 470 ThrAspGlyIleleuMetAspIleSerProIleGluPheSerAlaSerAlaSerleuGln 489  
Dh 4889 AAGAGAGGTAAACCCGTATAAATAGAGAAATAGACTTTAAGAGACATCTTATTGGACA 4948  
Oy 490 AspLysleuAsnleuValAlaLysAlaArgHisleuTyAspMetProAspArg 509  
Dh 4949 AATGAACAACCTCTTAACTAACAATAATGATGATCAATAAACATAGACAGATATTTT 5008  
Oy 510 ValIleAlaIleAsnHisAspAspGlyValAlaArgSerIleleuGlyArgIleSerAsp 529  
Dh 5009 TCTCTTTCTGATGCAATGATG-----GTAAAGCATTAACAGATTTTATTCAGAT 5062  
Oy 530 AlaValSerAlaValAlaArgAlaIleleuProAspGluSerGluAsnGlu----- 546  
Dh 5063 TATGTTATGCAATTGCTGAATAGATCTT-----GAACAACAATAATGAAGACAGATGCA 5119  
Oy 547 -----ValIleAspLeuProGluArgThrAlaLeu----- 556  
Dh 5120 ACAATTGATGTACTTCTCTTATGATTAACAAAGAGCTTCTTCTGTAAGATTAAT 5179

Oy 557 -----AlaAsnArgLysThrProAlaSpvalTyP----- 566  
Dh 5180 GTTGAAGAAATATCTGTAAGATTAATGTTATCTCCGTGAATTAAGCCCTGCTGAT 5239  
Oy 566 ----- 566  
Dh 5240 GGAGATCTTTTAATGTCACATCTCGACGCTTAATGAATGCTTAACCGCTTGGC 5299  
Oy 567 -----GlnSerLysValProleu 573  
Dh 5300 TATTTTAACCAAGTATACAGATACATCTCCCTACAGGAAAGAAATGATGATGATCA 5359  
Oy 574 TyValPheValAlaSerAspLysProArgAspGlnIleGlyLeuGlyTyPArgIlySer 593  
Dh 5360 CTGTGAAGGT-----CAAGACGCTCGAACAGGTGCATCAACAGGTGCTGTGCTTAC 5413  
Oy 594 AspThrGlyThrArgLeuValThrLys-----PheGluHisAsnleuIleAsnArg 610  
Dh 5414 TCACACATTTCTAAATTTGTTGCTTTCAGGAAGTATCTCAGAAAGAACTTATGGGAAAA 5473  
Oy 611 AspGlyTyArgGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysleu 630  
Dh 5474 ---GGTTAATTTTAAAGTATGTAAGCTTTATTTATGTAAGTCAATCTTCTGATCTT 5530  
Oy 631 TyAlaThrLysProleuSerHisProleuAsnAspGlnleuArgAlaThrleuGlyTyP 650  
Dh 5531 TCTTTTACCATCTCTGCTTTATGATACAGAC----- 5563  
Oy 651 GlnGlnGluValPheGluHisSerThrAsnGlyPheAspLeu----- 664  
Dh 5564 -----TTTGGCTTATGATTAACATTTATACCTGACAGATGAATGGATGAC 5611  
Oy 665 -----SerThrArgThrleuGlnHisGluIleSerArgSerIle 677  
Dh 5612 TTCCGTAAGAAACCTTATGAGATACCATACGTGATTTACACCTGTA----- 5659  
Oy 678 ILeGlnGlyGlyTyPAsnArgThrTyPSerLeuArgTyArgLeuAspLysLys 697  
Dh 5660 -----GGACAAATATTCATCTATCTTGTGGCGATGAGATGATCAATATCGT 5707  
Oy 698 ThrGlnAlaProProGluThrThrGlnAspLeuProValAspPheValAsnGlyLysPro 717  
Dh 5708 CTATATGATATTCATCTCATCTGACGCCACGCTCTATCTTGCATAT---CAAGCGAAAT 5764  
Oy 718 SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnleuVal 737  
Dh 5765 ATTTCTAGTGTAGTAAGTGTGTGTTTACTTTGATTTACA-----GACAGTCTGAG 5818  
Oy 738 AsnProMetArgGlyTyArgGlnArgTyPSerLeuGluValGlySerSerGlyLeuVal 757  
Dh 5819 AGACCATCTTAAGGGCATATTCGAAACAACTAATGTTGAATATGAGAGTGTGCTTTGGT 5878  
Oy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyPSerPheGlyAsp 777  
Dh 5879 GGTAAATGATACCTTCTCAAGCCCAATTGTGTAACGATCAAGAGATTTACTCAAT 5932  
Oy 778 AsnAlaTyGlySerAsnArgAlaHisGluMetThrGlyGlyIleGlnIleGlyTyPle 797  
Dh 5933 -----TCAAGAGTAAACCAACATTAATATCATTTGGCGTACAGGTGAGGCTGAGCT 5983  
Oy 798 TrpSerAspAsnPheAsnHisValProTyArgLeuArgPhePheAlaGlyIlyAspGln 817  
Dh 5984 TATTAAGATATGAATAAACCTGTGCGCAGATTTTGCAGCAATTTTATGTGGTATAGAT 6043  
Oy 818 SerIleArgGlyTyArgAlaHisAspSerLeuSerProIleSerAspLys---GlyTyPleu 836  
Dh 6044 AGTATTAAGAGATATGATACAGAGATCTTGGCACCACAAAGATCTCGCTTGGAGATGAA 6103  
Oy 837 ThrGlyGlyGlnValleuAlaValGlyThrAlaGlyTyArgTyPArgGluPheMetLysAsp 856  
Dh 6104 ATTTGGTGTAGTATGAGTGTCTTTTCTTAACSTAGATTAATTTTGGACATTTCCAGCCAGAG 6163



```

Oy 459 Thr---LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp--- 476
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1069 AACGGCTTGTAAGATCATTTATCGATTGAGAGTGGCGATATGGTGTATATCATATGAT 1128
Oy 477 -----IleSerProIleGluPheSerAlaSerAsnLeuIle----- 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1129 GTCATCATTTTCAGGAGAACCGCCAGCGATGATGATCTTATGAAAGGAGTATTGTTTA 1188
Oy 489 -----GlnAspLysLeuAsnLeuValAlaAlaLysAlaAlaArgHisLeuTyraAspMetPro 506
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1189 GGGCCCTAAGGATTAATACACTTGACCAACTGAGA----- 1224
Oy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1225 -----ATTCGCAAAATCTTTAAAGCGCT-----TTAGGATTC 1257
Oy 527 IleSerAspAlaValSerAlaValAlaAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1258 TTCTCTAAGTCAAAATTGAAAGAAAAAGGCTT-----AATAGCTCA 1299
Oy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1300 CTCATCGATTTA----- 1311
Oy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1312 -----TTAGTGAGCGTAGAAGAGGGCGCTACTGGGCGAG 1344
Oy 587 Ile-----GlyLeuGlyTrrpGlySerAspThrGlyThrArgLeuValThrLysPheGlu 604
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1345 TTGCAATTTGGGTAGCGATGCTCTTATGAGGGCTTATGCTT-----AAT 1392
Oy 605 HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAsp 624
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1393 GGGAGCGTGAGGAAGAAGACCTTTTGGCACAGGCGCAAGCATGAGCTTGTATGCTAAC 1452
Oy 625 -----LysLysGlyValLysLeuTyrAlaThrLysPro----- 635
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1453 ATCGGTACAGGGGGGCGGTATCTTATCCGGCGATGCAAAAGAGCGGGCGTATGTTT 1512
Oy 636 -----LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1513 GCGGGGAATTTGAGCTTGACTAATCCA----- 1539
Oy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeu--- 669
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1540 -----AGCATTTT-----GACAGCTGTATAGCTCTACATCAACCTTTAT 1581
Oy 670 ---GluHisGluIleSerArgSerIleIleGlnAsnGlyGly----- 682
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1582 GCGGATTCACGATTAAGCTACCAATACATCAACAAGCGGGGCTTGGGGTGATGTC 1641
Oy 683 -----TrrAsnArgThrTyr---SerLeuArgTyrArgLeuAspLysLeuLys 697
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1642 GGGCCCATCTCGGTAAATGAACCATGAGCTTAAGGCTTAATCACTTAATCAACAA 1701
Oy 698 ThrGlnAlaProProGluThrTrrpGlnAspLeuProValAspPheValAsnGly----- 715
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1702 CTCCTTGTTGTTAGCAGCCCTTATACAAACCGCTACTATTCCTGTTAATGAAGTGTT 1761
Oy 716 LysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAla----- 733
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1762 TCTCCAAAGCAATGTCTACCCCGCATCGGATATCAATCGCTTATCAGCGGTAA 1821
Oy 734 -----AspAsnLeuValAsnPro-----Met 740
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1822 ACCCCCTTACAACCTGAAGGCTTCTAGTCTGAGCGATCAACCACTTCACAGAAATA 1881
Oy 741 ArgGlyTyrArgGlnArg-----TyrSerLeuGluVal 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1882 AGAGTATTGGGATGAGGATTAACCATAGCGCTATCACAGCTTTTACCCCTTGATGTG 1941

```

```

Oy 752 Gly-----SerSerGlyLeuValSerAspAlaAsn 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1942 AGCTATGACACACCCATGATTTACTTCCCTAGAAATGGGTTATCTTACTG----- 1995
Oy 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp----- 777
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1996 -----TCTATCGACAGATGCTGGGCTTCCCAAGCTCTGGCACGCTCATATCTTGG 2046
Oy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2047 AACGGGTTAGCGGGAATGCTCGTAACACCAAGTTATAGTAAATTCGCGGCTTACAC 2106
Oy 792 -----IleGlnAlaGlyTyr 796
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2107 CATTTGCAAAATATTATTTGATAGATTTGATCGCTTCGCTTTAAACGCAAGAGGTTAT 2166
Oy 797 IleTrrSerAspAsnPhe---AsnHisValProTyrArgLeuArgPhePheAlaGly 815
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2167 ATCTTATAGTATATACACCGCATGATTTACTGCCCTTAACCTCCACCTTCTACATGGGGGG 2226
Oy 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyr 835
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2227 GTAACCAACGGTGAGGCTTTAGCAAGGATGCGTTACTCTTAAGATGAGTTGGCTTTC 2286
Oy 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLys 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2287 TGCGCTTGAGAGCGATGGGATTTTACCCTTCTACTGAAATGAGCTATGGGGTGCTAAG 2346
Oy 856 -----AspLeuArgLeuAlaValPheGlyAspIleGly----- 866
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2347 GCGGCTAAATAGCGCTTACGCTGTTTGTGACTTGGTTTCTTAACCTTAACCCCA 2406
Oy 867 -----Asn 867
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2407 ACTAGAGGAGTTTTCCTATTAACGCTCTGTTACGACACACCGAATTTTAAGATTTCG 2466
Oy 868 AlaTyrAspLysGlyPheThrAsnAspThr---LysIleGlyAlaGlyValGlyValArg 886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2467 GTATATGAGGCTGGGTTTGAAGAGCGACTTGGAGGCTTCCACAGCGCTTGCAAGTTGA 2526
Oy 887 TrrPalaSerProValGly 892
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2527 TGGATTTGCGCCATGGGG 2544

RESULT 20
US-10-335-977-157
Sequence 157, Application US/10335977
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-Dec-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.

```

```

1      REGISTRATION NUMBER: 36.207
2      REFERENCE/DOCKET NUMBER: GTN-018
3      TELECOMMUNICATION INFORMATION:
4      TELEPHONE: (617)227-7400
5      TELEFAX: (617)742-4214
6      INFORMATION FOR SEQ ID NO: 157:
7      SEQUENCE CHARACTERISTICS:
8      LENGTH: 2778 base pairs
9      TYPE: nucleic acid
10     STRANDEDNESS: double
11     TOPOLOGY: circular
12     MOLECULE TYPE: DNA (genomic)
13     HYPOTHETICAL: NO
14     ANTI-SENSE: NO
15     ORIGINAL SOURCE:
16     ORGANISM: Helicobacter pylori
17     FEATURE:
18     NAME/KEY: misc.feature
19     LOCATION: (B) LOCATION 1...2778
20     SEQUENCE DESCRIPTION: SEQ ID NO: 157:
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
10
```

```

Db 1405 -----TTA 1407
Oy 577 ValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTyrProGlySerAsp 594
Db 1408 GGAACCGTAGAAGAGGAGGACACCGGAGTTGCAATTCGAGTGGGCTATGGCTCTTAT 1467
Oy 595 ThrGlyThrArgLeuValThrLysPheGluHisLeuIleAsnArgAspGlyTyrGln 614
Db 1468 GGAGGGCGTCATGCTT-----AATGGAGCGTAGGGAAGAAGATCTTTTGGC 1515
Oy 615 AlaGlyAlaGluLeuArgLeuSerGluAsp-----LysLysGlyValLysLeuTyrAla 632
Db 1516 ACAGGGCAAGAGCATGAGCTTGTATGCTACATTCGCACAGGGCGGCGGATCTTATCCG 1575
Oy 633 ThrLysPro-----LeuSerHisProLeu 640
Db 1576 GGCATGCCAAAAGGGGGGGCGGTATGTTGCCGGGAATTGAGCTTGACTATATCCA--- 1632
Oy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnIleValPheGlyHisSerThrAsn 660
Db 1633 -----AGGATTTT-----GAC 1644
Oy 661 GlyPheAspLeuSerThrArgThrLeu-----GluHisGluIleSerArgSerIleIle 678
Db 1645 ACCTGGTATAGCTCTACGATCAATCTTTATGGGATTACAGATTAAGTACCAATACATC 1704
Oy 679 GlnAsnGlyGly-----TrpAsnArgThrTyr 687
Db 1705 CAACAAGGGGGGGCGCTTGGGGAATGTCGGCGCATCTGGGATATAGAACCCATGTG 1764
Oy 688 SerLeuArgTyrArgLeuAsp-----LysLeu-LysThrGlnAlaProProGluThr 705
Db 1765 ACCTTAGGGTATTAAGTGAATGTTACCAAACTCTGTGGTTGACAGCCCTTATACAA 1824
Oy 705 PcdInsPleuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaI 725
Db 1825 GCCTACTATTCCTGTGTATTAATGAAGTG--GCCTCTCCAAAGCAATGTTCACACCCGCA 1883
Oy 725 ValAlaValHisLys-----ThrValAlaAspAsnLeuValAs 738
Db 1884 GGTGATTATCAACCGCTTATCAAGCGGTAGACTCCATTGGTCCCGTAAGCGTCTAG 1943
Oy 738 nPro-----MetArgGlyTyrArgGlnArg----- 746
Db 1944 TCCTGAGGAGATCACCACTTCAACAGAATAAAGGTATTGGAGGATTACACAC 2003
Oy 747 -----TyrSerLeuGlnValGly----- 752
Db 2004 GCCTATCACCACTTCTTACCCCTTGATGTAGCTATGACACACCGCATATTATT 2063
Oy 753 -----SerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerG 771
Db 2064 CCCTAGAATAATGGGCTTATCTTAGT-----TCTATACGACAAATGCTCG 2108
Oy 771 ValTyrSerPheGlyAsp-----AsnAlaTyrGlySerAsnArgAlaHisG 787
Db 2109 TTGGCAACGCTCGCACCTCAATCTTGAACGGGTTAGCGGGAATGTCGTAACAC 2168
Oy 787 nMetThrGlyGly----- 791
Db 2169 CAAAGTTATATGTAATATTCGCCGCTTACCAACATTTGCAAAATATTATGATAGATT 2228
Oy 792 -----IleGlnAlaGlyTyrIleTyrSerAspAsnPhe-----AsnHisVal 805
Db 2229 GATCCGCTGTTTAAACGCAAGGAGGCTATATCTTATAGTATACACCGCATTAATCT 2288
Oy 805 LProTyrArgLeuArgPhePheAlaGlyLysArgGlnSerIleArgGlyTyrAlaHisAs 825
Db 2289 GCCCTTAAATCACTTCTTACATGGGGGCGTAACACAGGTGAGAGGCTTTAGAACGG 2348
Oy 825 pSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValAlaI 845
Db 2349 CTCATACACACTTAAGATAGTTGGCTTGGCTTGGAGGATGGGATTTTACCGC 2408

```

```

Oy 845 yThrAlaGluTyrAsnTyrGluPheMetLys-----AspLeuArgLeuAlaValPheG 863
Db 2409 TTCCTACTGATTAAGCTATAGGGGTGTAAAGCGGCTTAAATGCGTTTAGCGGTTT 2468
Oy 863 yAspIleGly----- 866
Db 2469 TGACCTTGGTTTCTTAACCTTAAACCCCACTAGAGGGGATTTCTCTATTAACGCTC 2528
Oy 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspH 877
Db 2529 CACACAGAGCGGAATTTAAAGATTATGCGCTTGAGGGCGGCTTGAAAGGGCGAC 2588
Oy 877 r---LysIleGlyAlaGlyValGlyValArgTyrPheAlaSerProValGly 892
Db 2589 TTGGAGGCGCTTCTACAGGCTTACACAGTTAATGATTTTGGCCCATGGGG 2637

RESULT 21
US-10-419-128-4442
; Sequence 4442, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4442
; LENGTH: 2466
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-4442

Alignment Scores:
Pred. No.: 4.3e-05 Length: 2466
Score: 186.00 Matches: 175
Percent Similarity: 33.74% Conservative: 132
Best Local Similarity: 19.23% Mismatches: 276
Query Match: 3.93% Indels: 330
DB: 8 Gaps: 42

US-09-914-168-2 (1-919) x US-10-419-128-4442 (1-2466)
Oy 223 SerAlaMetAspLeuAsnGlySerIlePro-ArgLeuArgGlnThrAlaLeuValAlaI 242
Db 5 TCTGTAGTCAATGTGAATTCGAATCTGCCGCCCTTTGGCGGACATTTCTATTTGCACT 64
Oy 242 aArgAlaValAlaGlyTyrTyrAspPheLysPheSerIleIleArgAsnSerIleGlyL 262
Db 65 TGGAAAAAAGAGCACTCCAGAA---ACGCTTCTGCTACCCCGCGTGTGTCAGCCCTGA 121
Oy 262 LAsp-ValIleIleHisAspLeuGlyLuproValTyrIleAspTyrArgAlaValAlaI 282
Db 122 TCAATCCCGAGGTTTCAAC-----GCCGAGTCTTCACTGTTCCGATA 163
Oy 282 aLArgGlyGlnGluAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu 302
Db 164 TCGGGTCAATGGCTTCGACAGCGGTGTCCGCGGTAGCGTTCGCCGCCCTCCCTGA 223
Oy 302 euIleGlyAsp---ValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 321
Db 224 ACCTCGGACAGACATGACGACACAGCCCTGATCCAGGCAACCCGCTTCGTCACAGA 283
Oy 321 snAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 341
Db 284 CC-----GGCTTCTTCCAGGACATTCACAGCTCGGCGTGTATGGCAACGCTGC 328

```

Oy 341 leLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgP 361  
 Db 329 TG----- 330  
 Oy 361 heAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 381  
 Db 331 -----GTAAGTCACTAGTCGAGCGCCCTCATCTCCAGCATTCAGATCGAGGGCA 382  
 Oy 381 sPLysLeuProValLysArgGlnLeuGlnGlnLeuLeuThrValAsnMet-----G 399  
 Db 389 ACAAGCGCATTTCCAAAGCAAGCACTGCTCAAGCGCTCAAGACAGTCCGGCTCGCGCAAG 442  
 Oy 399 lyGlnAlaTyrTrsLeuGlnAlaValAlaArgAlaLeuSerAspAspLeu----- 414  
 Db 443 GCGAGATCTTCAGCGCGCCACCCTCGAAGCGCGTGCAGACAGCTGCACACCGCCAGTACG 502  
 Oy 415 --IleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPheProGlnArgGln 434  
 Db 503 TGGCCCGAGCGCGCTACTCGCGCGAGATCAACGCGAGATGATCCCGCGCGCGCAAC 562  
 Oy 434 IuIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThrGlnProAlaG 454  
 Db 563 GC----- 564  
 Oy 454 InValAspGlnSerThrLeuGlnProValIleGlnThrValGlnLeuThrAspGlyIle 474  
 Db 565 -----GTCGGCTGAAGATCAACATCAACAGCAAGCAAGCAAGCAAGCAAGCAAG 598  
 Oy 474 euMetAspIleSerProIleGlnPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnL 494  
 Db 599 TCGCCCGCATTTCCACATCAACGCTGTCGCGCAACCGCTCTTCTCCGAAGA----- 651  
 Oy 494 euValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValIleAla---I 513  
 Db 652 -----GACCTGACCGACCTGTTCGAACTGAAGACCAACCTGCTGCTCTCT 700  
 Oy 513 leAsnHisAspAspGlyValAsnArgSerIleLeu----- 524  
 Db 701 TCAGAAACGACGACAGTACGCGCGAAACCTTCGCGCGACCTCGACCGCTCGCT 760  
 Oy 525 -----GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIle 539  
 Db 761 CCTATTACTGACCGCGCTTACATCAACATGATATCGCTTCACCGCGAGTTCATCA 820  
 Oy 539 euProAspGlnSerGlnAsnGlnVal-----IleAspLeuProGlnArgThrAlaL 556  
 Db 821 CCGCGGACAAAGACGACGCTTACATCAACGCTCAACAGAGGAGAGACATCAACCA 880  
 Oy 556 euAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro----- 572  
 Db 881 TCCGCGACGTGAAGCTGACCGCGGACCTG-----AAGGTCCCGAAGAGAGAG 928  
 Oy 573 -----LeuTyrValPheValAlaSerAspLysProArgAspGlnIleGlyL 589  
 Db 929 TGAAGCCCTGCTGCTGCTA-----CAGAGGCGCAGGTGTT-C 966  
 Oy 589 euGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleA 609  
 Db 967 TCGCGCAAGCGTGAATGAC-----CACCACCTC----- 992  
 Oy 609 snArgAspGlyTyr-----GlnAlaGlyAlaGlnLeuArgLeuSerGlnLysP 625  
 Db 993 --CGACCTAATCAACCGCTCGCTGCGAAGCGGCTACACCTTCGCGCAACGCTAACGGC 1050  
 Oy 625 yAluSglnValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuA 645  
 Db 1051 GTCCCGGAGAC--CCCATGACAGCAAGCAAGCGCTCGTGAACCT--TCGTCGTGCAACCC 1106  
 Oy 645 rGAlaThrLeuGlyTyrGlnGlnGlnValPheGlnHisSerThrAsnGlyPheAspLys 665  
 Db 1107 GGGCAAGCGC--GCCTACGTCGAACCGCATC-----AATTTCGCGGCA 1147

Oy 665 eThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGln--AsnGlyGlyTyrP 684  
 Db 1148 ACACCAAGACGAGACAGCAAGTGTGCGCGCGAAATCGGCAGATGTGAAGCGGCTGGG 1207  
 Oy 684 snArgThrTyrSerLeu-----ArgTyrArgLeuAspLysLeu----- 696  
 Db 1208 CCTCCACTACTGTGATTCACACCAAGTCCAAAGCGCGCTGAGCGCTCGCTACTTCAAG 1267  
 Oy 697 -----LysThrGlnAlaProProGlnThrThrGlnAspLeuProValAspPhe- 712  
 Db 1268 AAGTCAACCTGAGACACCGCGCGGTTCCTCCGGCACCGAGACAGGTGACAGTCAACTCA 1327  
 Oy 713 --ValAsnGlyLysProSerGlnGlu-----AlaLeuLeuAlaG 725  
 Db 1328 GCGTCGAAGACCAACCGTCCGCTGATTCACCGACCGCGTGGCTTCCCGAGCGCGC 1387  
 Oy 725 lyValAlaValHisLysThrValAlaAspAsn----- 735  
 Db 1388 GCGTATCTCGCGCGCTGATTCAGCAAGAACACTTCTCGGTACCGGCAACAGTCA 1447  
 Oy 736 -----LeuValAsnPro- 739  
 Db 1448 GCATCGCGCTGACCGGACGAATACAGACCGCTACAACTTCGCTTCGTCGACCCCT 1507  
 Oy 740 -----MetArgGlyTyrArgGlnArgTyr----- 747  
 Db 1508 ACTGACCGCTGACGCGCGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1567  
 Oy 748 -----SerLeuGlnValGlySerSerGlyLeuValSerAspA 760  
 Db 1568 ACAGACTCGACGTGACGTCGCGCAGTACTCGGTGACAGCGCTGGC-----G 1615  
 Oy 760 laAsnMetAlaIle----- 764  
 Db 1616 CCGCATGAGCATCGGCTACCGCATGACGAGACCTGCGCTGACTGAGCTGAGCG 1675  
 Oy 765 -----AlaArgAlaGlyLysSerGlyValTyrSerPhe----- 775  
 Db 1676 TGCAGCGCACACGATGACACCGCGCTGCTGACCGCTGACAGAGATCTGACGCTTCG 1735  
 Oy 776 -----GlyAspAsn-----AlaTyrGly----- 781  
 Db 1736 ACNAGAGAGCGCACACTTCCACCACTTCAGAGCTTCGCTGCTGCTGCTGCTGCTGCTGCT 1795  
 Oy 782 -----SerAsnArgAlaHisGlnMetThrGlyIleGlnAla----- 794  
 Db 1796 TGAACAAGCGCTGCTGCGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1855  
 Oy 795 -----GlyTyrIleTyr----- 798  
 Db 1856 TGGCGGCGACGACCTGTGCTTACAGATGAGTACGAGCGCGCGAGTCTGCGCGC 1915  
 Oy 799 --SerAspAsnPhe----- 802  
 Db 1916 TGACCGCACTACACCATGCGCTTCCACACGAGCTGCGCTTACGCGCGCTGCGCGC 1975  
 Oy 803 -----AsnHisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArg 821  
 Db 1976 CCAACGAGCGCTGCGCTTACGAGACACTACGCGCGCGCTTCAACTGCGTGGG 2035  
 Oy 821 lyTyrAlaHisAspSerSerProIleSer----- 831  
 Db 2036 GCTTCAGAGACAGACGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2095  
 Oy 832 -----AspLysGlyTyrLeuThr----- 837  
 Db 2096 AGATCATCTCTGCGCTGATGAGCGTGTGCTACACGAGCGCGAGACGAGATCCGAG 2155  
 Oy 838 -----GlyGlyGlnValLeuAlaValGlyThrAlaGlu-----TyrAsnTyrGlnPhe 854  
 Db 2156 CTTTCGCTGCTGATCTGATCTGATCAAGCGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCT 2215  
 Oy 854 eLysAsp-----LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrTrsP- 870



	Db	2216	TCAAGACACCAGCCCACTTTCCTGATCCGTCTGTTCTGGGACTGGGTAGCATTCTGCATA	2275
	Oy	871	-----LysgLyPhe <sup>Thr</sup> Asnasp <sup>Thr</sup> LysIle-----	879
	Db	2276	CCGACTGCCGACCAAGAACCCACCACTGCACGCGCATCAAGACCGCAACACTCGCCA	2335
	Oy	880	--glYalaglyAlglyValargTrpAlaserProvalIglnValArgValaspyla	899
	Db	2336	GTTCCGCGGGCGCTGCGCTACCTGCATCACCGCGCTGGAGCCCGCTGAAGCTTCAAGCTTG	2395
	Oy	899	IarhrglyValLysgLnlglyAsn	907
	Db	2396	CGAGCCGATCAAGAAGCCGGAAC	2421

```

Qy 684 snatgthrtyrserleu-----Argrtyrargleuasplysleu----- 696
Db 1208 CHTCCACCTGATCGTACGACCAAGTCCAGGCCGCTCGAGCGCTCGGCTACTTCAAG 1267
Qy 697 -----LysThrGlnAlaProProGlnThrTrpGlnAspLeuProValAspPhe- 712
Db 1268 AAGTCAACGTGACAGACCCGCGGTCCCGGACCGAGACAGTACAGTCAACTCAACA 1327
Qy 713 --ValAsnGlyLysProSerLeuIncln-----AlaLeuLeuAlaG 725
Db 1328 GCGTCGAAGAACCAACCGTCCGCTCGATCACCAGCGCTGGCTCGCCAGAGCGCG 1387
Qy 725 LysAlaAlaValAlaHisLysThrValAlaAspAsn----- 735
Db 1388 GCGTGATCTCGGGGGGCTCGATCAACCAACAATTCTCGGTACCGCAACAAGTCA 1447
Qy 736 -----LeuValAsnPro- 739
Db 1448 GCATCGGCTGACCCGAGCGAATACCAAGCCCGCTACAACCTCGGCTTCGTCGACCCCT 1507
Qy 740 -----MetArgGlyTyrArgGlnArgTyr----- 747
Db 1508 ACTGACCGCTGACAGCGCGTCAACCGCTCGCTACAACGCCCTTCAACCGAAGACCGACTACG 1567
Qy 748 -----SerLeuGlnValGlySerSerGlyLeuValSerAspA 760
Db 1568 ACGAATCGACCTGACAGCTCGCGACCTACTCGTGAACAGCCTGGGCG-----G 1615
Qy 760 LAsnMetAlaLe-----MetArgGlyTyrArgGlnArgTyr----- 764
Db 1616 CCGGCATGACATCGGCTACCGGATCAGGAGACCTCGCGCTGACCTATGGCTGAGCG 1675
Qy 765 -----AlaArgAlaGlyLysSerGlyValTyrSerPhe----- 775
Db 1676 TCGAGCGCAGCAGATCGACACCGCGCGCTACACCGTTCAGCAGATTCAGCTTCCTCG 1735
Qy 776 -----GlyAspAsn-----AlaTyrGly----- 781
Db 1736 ACAAGAGAGCGACACACTTACCAACTTCAGAGCTTCATCGGCTGCGAATCGAGCC 1795
Qy 782 -----SerAsnArgAlaHisGlnMetThrGlyGlyLeuGlnAla----- 794
Db 1796 TGAACAAGCGCTGCTGGCCACCGCTGCTACGCGAGACCGCTGGAACCAACCC 1855
Qy 795 -----GlyTyrIleTrp----- 798
Db 1856 TCGCGGCGACCGACCTGTCTCTACAGATGCTACGCGGCGCAGGTCTTGGCCCGC 1915
Qy 799 --SerAspAsnPhe-----GlyTyrIleTrp----- 802
Db 1916 TGACCGGACACTACACCATCGCTTCCACACGAGCTGGGCTATGGCGAGCGCTACGGAT 1975
Qy 803 -----AsnHisValProTyrArgLeuArgPhePheAlaGlyAspGlnIleArg 821
Db 1976 CCACCGAGCGCGCTCGCTTCTACGAGACTACTAGCGCGGCGCTTCACTCGTGGCGG 2035
Qy 821 LysTrpAlaHisAspSerLeuSerProIleSer----- 831
Db 2036 GCTTCAAGGACAGCAGCTGGGCGCTGAGTACCCCTCTAGGGCTTACAAAGATGGCA 2095
Qy 832 -----AspLysGlyTyrIleThr----- 837
Db 2096 AGATCATTCCTGGCGCTGATGAGCGTGTCTGCTACACCGAGCCGAGACAGATCCGAG 2155
Qy 838 -----GlyGlyGlnValLeuAlaValGlyThrAlaGlu-----TyrAsnTyrGluPhe 854
Db 2156 CTTTGGTGGTGAACATCTGATCACCGCGCGCGAGAGCTCTTCCCGCTGGCGTTGCG 2215
Qy 854 eLysAsp-----LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp- 870
Db 2216 TCMAAGGACACCGCGCAGTGGCTACCGCTGTTCTGGGAGCGGTGAAGCACTTCGATA 2275

```

```

Qy 871 -----LysGlyPheThrAsnAspThrLysIle----- 879
Db 2276 CCGACTGCCCGACACCAACCAACCACTGCGAGCGCATCAAGACCGACACTCGCCA 2335
Qy 880 --GlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValA 899
Db 2336 GTTCCGTTGGGCGTCCGCTACCTGATCACCGCGCTGGCGCGCTGAGCTTCAAGCTGG 2395
Qy 899 LathrGlyValLysGluGluGluAsn 907
Db 2396 CGACGCCGATCAAGACCGGAGAAC 2421

RESULT 23
US-10-419-128-14662
Sequence 14662, Application US/10419128
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NOCUTIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/10/419,128
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 14662
LENGTH: 516
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-10-419-128-14662

Alignment Scores:
Pred. No.: 4,95e-05 Length: 516
Score: 174.50 Matches: 49
Percent Similarity: 46.28% Conservative: 38
Best Local Similarity: 26.06% Mismatches: 62
Query Match: 3.69% Indels: 39
DB: 8 Gaps: 6

US-09-914-168-2 (1-919) x US-10-419-128-14662 (1-516)
Qy 230 SerIleProArgLeuArgGlnThrAlaLeuVal-----AlaAlaArgAlaVal 245
Db 23 GCGTTCACAGCTTCCGTGCGCAATGCCGAGCGCGAGCGGAAAGCGCGCCAGCGGCTC 82
Qy 246 GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyValAspValIle 265
Db 83 GCCTACTTCCAG-----GCGCAATCGACACCGAG 112
Qy 266 IleHisAsp-----LeuGlyGluProValTyr 274
Db 113 GTGAGAGACGCGCAAGCCGCGAAGCTCAGCTCAAGTACTGCCCCGCGAGCCGCTGCGC 172
Qy 275 IleAspTyrArgAlaValArgGlyGluGlyAlaAspAspLysAlaPheThr 294
Db 173 CTGGCCAGGTGAACATCCAGGTGCTCGGAGCGCGCAGCGCTGAGAGACTTCCGCTTG 232
Qy 295 ValAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThr 314
Db 233 CCGTGGGCGCAAG--CAACTGAAGCGGCGCGAAGCTGAACGAGGTGTCTACGAGGAC 289
Qy 315 LysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTrpLeu 334
Db 290 GCCAAGCGGCGATCGCAAGCAAGCCTCGGCTACGAGCTTCTTCCAGGGCGGCTTACG 349
Qy 335 AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp 354
Db 350 ACCCAGCGCTGACATCGATCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 409

```



Db 781 CGACAACTTCCACCAATGTTCAAGGCTTCATCGGCTGGCTGGAAATCGAACCCCTGAACAGGG 722  
 Oy 676 rlelleclnanslgllytprasnargthtyrserleuargtyrargleuasplysle 696  
 Db 721 CGTGCTGGCCACCGCGGTGCTACCTCGACAGACCTGACC-----CT 683  
 Oy 696 ulysthrclnalnarproprnogluphtprglnarpleurprovalasprhevalaanglyly 716  
 Db 682 GGAAGCCACCCCTGCCGGGACGACGACCTGTCTGTACAAAGATCGACATAC---CGCGGGCA 626  
 Oy 716 sprosercnglnalaleuleualaglyvalalavalahlsystrhvalalaaasprnle 736  
 Db 625 G-----GTCTTCGCCCCGGGACCGACCAAC-- 601  
 Oy 736 uValasnrpromeTarGlytyrArglnArgprysrleuGlnuValGlyserSerGlyLe 756  
 Db 600 -----TACACCATCGGCTTCACACCCGACGCTGGGC----- 571  
 Oy 756 uValserAspAlaasmetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGl 776  
 Db 570 -----TATGG 566  
 Oy 776 yAspAsnalArgGlyGlySerAsnArgAlaHisGlnmetThrGlyIleGlnAlaGlyTy 796  
 Db 565 CGAC---GGCTACGGATCCACCC----- 547  
 Oy 796 rlelrrprserAspAsnrPheAsnHisValrProtyrArgleuArgPhePheAlaGlyLyAs 816  
 Db 546 -----GACGGCTGGCGGCTTACGAGACACTACTAGCGCGGGGCTT 506  
 Oy 816 rclnserlleargGlytyrAlaHisAsrserleuSerProIleSer----- 831  
 Db 505 CAACTCGTCGGCGGCTTCAGAGACAGCAGCTGGGGGCTCTGTAGACCCCTTAGGGC 446  
 Oy 832 -----AsplysGlytyrLeuThr----- 837  
 Db 445 TTACAAAGATGGCAGATCATCTTCCTGGGCTGATGAGCGCTGGCTACACGACCCGGA 386  
 Oy 838 -----GlyGlyGlnValIleuAlaValGlyThrAlaGlu-----ty 849  
 Db 385 CCAGATCCCGCAAGCCTTCCTGGCTGTACATCTCGATACCGGGGGCGCCAGCTGCTGT 326  
 Oy 849 rAsnTyrgluphemeTyrAsp-----LeuArgLeuAlaValPheGlyAspIleGl 866  
 Db 325 CCGCGTCGCGCTGTGTCAAGGACCCAGCGCCAGTGGCTTACCGTGTCTTGTGGGACGTGG 266  
 Oy 866 yAsnalArgAsp-----LysGlyPheThrAsnAspThrLysIle----- 879  
 Db 265 TACGACCTTCGATACCGACGACGTCGCCGCAAGACCCACCACTGGACGCGCATTCAGAC 206  
 Oy 880 -----GlyAlaGlyValGlyValIleGlyArgTrpAlaSerProValGlyLysIle 894  
 Db 205 CGAACAACTTCGCGCATGTCCTGGGCTGGCTGGCTGACCTGGATACCGCGCTGGCCCGCT 146  
 Oy 894 lArgValAspValAlaThrGlyValLysGlnGlnGlyAsn 907  
 Db 145 GACGTTGACGCTGGCGACGCGCATCAAGAGCCGCAAC 106  
 RESULT 26  
 US-10-366-683-4165/c  
 ; Sequence 4165, Application US/10366683  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rubenfeld, Marc J.  
 ; APPLICANT: Nollung, Jork  
 ; APPLICANT: Deloughery, Craig  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: PATHO3-04  
 ; CURRENT APPLICATION NUMBER: US/10/366,683  
 ; CURRENT FILING DATE: 2003-02-13  
 ; PRIOR APPLICATION NUMBER: 09/252,991





Db 305 GAAGATTGACTGCGACGACCAACCTATTGA-----GATCCCAATCCAAATTTAAACCA 358  
 Qy 401 AlATyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420  
 Db 359 ATGTACATATATATAATTAATACATATACATATACATATACCTTATA----- 403  
 Qy 421 AsnMetValAsnThrGluIle--ValPheProGluArgGluGlnIleGlnAsnAspGln 439  
 Db 404 -----GTGAAGCTGCATATCTCCCTCTTCTCTTAGAAGAAATTCATC----- 445  
 Qy 440 ValSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspLeuSerThr 459  
 Db 446 -----TCT 448  
 Qy 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLeuSer--- 478  
 Db 449 CTAAAGTAAATCCCTG-----CTCTTGCTGTTAATATCTTAT 484  
 Qy 479 ProIleGluPheSerAlaSerAsn-----LeuIleGlnAsp----- 490  
 Db 485 CCTGTGTAAGTAGCCTCATGCAATATATATACATATATATGATATCCATTAGCCAAATTA 544  
 Qy 491 ---LysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArg 509  
 Db 545 AAAAATCTCAATAC-ATAATTATATCGATGATATACATTTGATATAT-----GATTTTA 597  
 Qy 510 ValLeu-----AlaIleAsnHisAspAspGly 518  
 Db 598 GTAATTAACACAGACCGCTTACTCATATGCGTAAAGTACACCGTCAACATATCTCAT 657  
 Qy 519 ValAsnArgSerIleLeuGluYArgIleSerAspAlaValSerAlaValAlaArgAlaIle 538  
 Db 658 ATTTCACAAATATCTCTTAATTAATTT-----GCTACATGAGAAAGAAAGAGGTA--- 708  
 Qy 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558  
 Db 709 -----TGGAAACATGACCTCTTCTGATTTCTTA-TCGAACACGGCTTCAACAAAC 755  
 Qy 559 ArgLysThrProAlaAspValTyr-----GlnSerLysLysValPro--- 572  
 Db 756 AGG-----CCTTTTCAGTCTATTAACCTCTCAATCCAAAGATCAAAAAGAAACAATG 809  
 Qy 573 -----LeuTyr-ValPheValAlaSerAsp---LysProArgAspGlyGlnIleGlyLe 589  
 Db 810 TAACACCCCTATAGAACTTTGTGCAACAGACCCCTCAAGACCTATTAGTGGGCTT 869  
 Qy 589 uGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAs 609  
 Db 870 ACAATACTCTTCTGATCAAGATGATTGTCACGTGGGACTTGGGAACAT----- 918  
 Qy 609 nArGAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLys 629  
 Db 919 -CGAATGTTTTTGTAAATGAGAACTTTTGTGATA----- 954  
 Qy 629 sLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuG 649  
 Db 955 -----ACAGCACCAATAGT-----CGAGATGATCAAAAATATTATGCAAA 995  
 Qy 649 yTyrGlnGlnGluValPheGluHisSerThrAsnGlyPheAspLeuSerThrArgThrIle 669  
 Db 996 CTTCCAAAACCAACCGCTTTGGCCGTCCAATATCAATCATTA---ATTAGGAACACACACT 1052  
 Qy 669 uGluHisGluIleSerArgSerIleIleGlnAsn-----GlyGlyTyrAs 684  
 Db 1053 TAAAAAAGAAATATACAAAAGTTACAAACACACACTGATCTATTGCTTTAGCAATTGA 1112  
 Qy 684 nArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluLuth 704  
 Db 1113 ACGACAATTTAATAGACGT----- 1131  
 Qy 704 rTyr-----GlnAspLeuProValAsp-----PheValAsnGlyLysProSerG 719  
 Db 1132 -TGGTTTGGTACTAGACACTTTTCAGTTGATACAGATTATATGATGATGAGATTCTAT 1190

Qy 719 nGluAlaLeu-----LeuAlaGlyVal-----AlaValHisLysThrValAlaAspAs 735  
 Db 1191 AAAAAAATATTACTCTTTTGGCATCCCTTATCATATACAGGAGATGTTCTTAAGA 1250  
 Qy 735 nLeuValAsnProMetArgGly----- 742  
 Db 1251 TCCTCTTATCTTATCCATTCACAAAGAAACAAAGCTACCTTAATGTTACTCTTATATGTTA 1310  
 Qy 743 -TyrArgGlnArg-----TyrSerLeuGluVal 751  
 Db 1311 ATATAAAAAAGATTTTTCCTTACCTTACGTAGTCGGTGTGATTTTATGCTTTATACATGAGCT 1370  
 Qy 751 lGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerG 771  
 Db 1371 TCTTAAACAGGAAACCTATCTTGCTTACAAA----- 1404  
 Qy 771 yValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyG 791  
 Db 1404 ----- 1404  
 Qy 791 yIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArg 811  
 Db 1405 -ATACCAATAGGTTCCTCTCAGGAAAGATATAGAAAATATCTCTGCAATACTAGGTT 1463  
 Qy 811 ePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIle 831  
 Db 1464 TTATGCTGGGGTGGTGGTGGTGAAGAGGTATGACTATCATCTTATGGACCAAAA 1523  
 Qy 831 rAspLysGlyTyrLeuThrGlyGlyValLeuAlaValAlaGlyThrAlaGluTyr-Asn 851  
 Db 1524 TAAATATGCGGATGCTATGAGAGA-----CTTCTTTTCAACATATAGTTGTAAT 1577  
 Qy 851 yrGlu-----PheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAla 869  
 Db 1578 ACGATTAAAAATTAACAGAAATCCATGTCATGTCCTCAATTTATGTCAGTGGGCAATATT 1637  
 Qy 869 yrAspLysGlyPheThr-----AsnAspThrLysIleGlyAlaGlyValGly 885  
 Db 1638 ACGAAAAAATAATTTCTGCTACTTAAAAAATCANTAT-TATGGGGGCTAGGCTGGGGC 1696  
 Qy 885 aLArgTyrAlaSerProValAlaGlyGlnValArgValAspValAlaThrGlyValLysGlu 905  
 Db 1697 TACGATATATATCAAGTTTGGCCCATACGTTTATGATATACCACTCCACTTCAAGATA 1756  
 Qy 905 lUGLYAsn-----ProIleLysLeuHisPhePheIleGlyThrProPhe 919  
 Db 1757 GAAGCATATATAAACACTTTCACCTTTATATTAGTATGCGCAAGCACTTC 1806

RESULT 29  
 US-10-369-493-42248/c  
 : Sequence 42248, Application US/10369493  
 : GENERAL INFORMATION:  
 : APPLICANT: Cao, Yongwei  
 : APPLICANT: Hinkle, Gregory J.  
 : APPLICANT: Slater, Steven C.  
 : APPLICANT: Goldman, Barry S.  
 : APPLICANT: Chen, Xianfeng  
 : TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 : FILE REFERENCE: 38-10(52052)B  
 : CURRENT APPLICATION NUMBER: US/10/369,493  
 : PRIOR FILING DATE: 2003-02-28  
 : PRIOR APPLICATION NUMBER: US 60/360,039  
 : NUMBER OF SEQ ID NOS: 47374  
 : SEQ ID NO 42248  
 : LENGTH: 1890  
 : TYPE: DNA  
 : ORGANISM: Halobacterium sp. NRC-1  
 : US-10-369-493-42248

Alignment Scores:

```

Pred. No.: 0.00718 Length: 1890
Score: 157.00 Matches: 99
Percent Similarity: 40.00% Conservative: 51
Best Local Similarity: 26.40% Mismatches: 133
Query Match: 3.32% Indels: 92
DB: 8 Gaps: 20

US-09-914-168-2 (1-919) x US-10-369-493-42248 (1-1890)
Oy 248 TyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHis 267
    :::::::::::
    1131 TTTCACCTTCATGCCAGGCTCAGCGCGCTCAGCTCACCACACACATGTCCTC----- 1078
Oy 268 AspLeu---GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGly 286
    :::::::::::
    1077 GACCTCCCGGAGACAGCCACCCCTGGACGGCGCGGAGCGCCAGCGCCACCCCTCGTGGG 1018
Oy 287 AlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal 306
    :::::::::::
    1017 GTT---GACGTTCTTCTGGGCTCTGCGCGAGAGAGTCTTCACCTCGACCTGGCGCTGACCTG 961
Oy 307 PheHisIleGlyLysTyrGluThrLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
    :::::::::::
    960 TGGCATTCTCTGGACCCCGCCGACGAGATCAGCTCGATGCTCCCTTTGCGAGAGCC 901
Oy 327 GlyTyrPheAspGly-----ArgTyrPheAspArgSerValAspValIleLeuProAsp 344
    :::::::::::
    900 GCGCTCTCGAGGCGCTCGCTCGCGCGCGCGAGCGCTGCGCTGATGAGTCTCGGAT 841
Oy 345 -----AsnThrAlaAspValSerLeu---IleTyrAspThrGlyThrGlyThr 359
    :::::::::::
    840 GCTCTCGAGAGCTGCGACGCGCTGATGCTGTTGAGAGTGTACGCGCGCGAGCGGCTGCG 761
Oy 360 ArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGluLeuThrThrAsp 379
    :::::::::::
    780 CGTGACGACGACGAGTT-----GAC 760
Oy 380 ProAspLysLeuProValLys-----ArgGluLeu 389
    :::::::::::
    759 CGTCGCTCTCTCGCGCTCGACAGCTCGATTGCTCTCGCGCGCGCTCGCTGAGTGC 700
Oy 390 LeuGluGluLeuLeuThrVal---AsnMetGlyGluAlaThrLysAsnLeuGluAlaArg 408
    :::::::::::
    699 CTGAGGCGCTGGCGGCTCTCCCGCAGCGAGTGCATCCGCTGCTTCCGAGTTGTCGC 640
Oy 409 AlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleVal 428
    :::::::::::
    639 GAGGTGCTCATATGCGCGTGCCTCA-----GTCGTCGCCGACGAGTGCCTT 592
Oy 429 PheProGluArg-----GluGlnIleGluAsnAspGluVal 440
    :::::::::::
    591 GTCCCTCTCTGCTGCTGCGACCTCTGAGACCGCGCGCGCGAGTCCGAGATCGACAGTGC 532
Oy 441 SerPheIleGluInsSerSerSerArgThrGluProIleGluValAlaAspGluSerThrLeu 460
    :::::::::::
    531 -----GAACTGCCCGCCACCGAGTCTGAGCCACCAACAGCGTCTG 493
Oy 461 GluProVal---IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479
    :::::::::::
    492 GTCCGCGCTCTCTGCGAGCGCGTACCGCATCGACCGCGC----- 454
Oy 480 IleGluPheSerAlaSerAsnLeuIleGluAspLysLeuAsnLeuValAlaAlaLysAla 499
    :::::::::::
    453 -----CCTCGCGCTCTGAGCATGCGCTCGACGCTC-----GAAAGCC 418
Oy 500 ArgHisLeuTyrAspMetProAspAspArgValLeuAlaIle-----Asn 514
    :::::::::::
    417 CGCGATCTCGCGCGCTCTGCTGCTGCGCTGCGCTGCGCTGAGTACGCGCGGAC 358
Oy 515 HisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaVal 534
    :::::::::::
    357 CGTGATACGCG-----TTTCTGACGCTCTCGCGCGAGTACTCTCGCGCTC 310

```

```

Oy 535 AlaArgAlaIleLeuProAspLysSerGluAsnGluValIleAspLeuProGluArgThr 554
    :::::::::::
    309 GCGTTGATCTCTGGAGATCGCGCGCA-----GATCTCTCGCGCTGTA 262
Oy 555 AlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyr 574
    :::::::::::
    261 CTCGTC-----GTCGCGGAGGCAACCGTGTACTGCTCTC 226
Oy 575 ValPheValAlaSerAspLysProAspArgAspGlyGlnIleGlyLeu 589
    :::::::::::
    225 GCCCATGTGCGCTTGATGCA---CCGATGTGCTGATCGGATTT 184

RESULT 30
US-60-453-134-135
; Sequence 135, Application us/60453134
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
; FILE REFERENCE: 28335/3815
; CURRENT APPLICATION NUMBER: US/60/453,134
; CURRENT FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 588
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 135
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: H. Influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (553)-(553)
; OTHER INFORMATION: n = a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (2238)-(2238)
; OTHER INFORMATION: n = a, c, g, or t
US-60-453-134-135

Alignment Scores:
Pred. No.: 0.0167 Length: 2349
Score: 154.00 Matches: 135
Percent Similarity: 31.37% Conservative: 94
Best Local Similarity: 18.49% Mismatches: 239
Query Match: 3.26% Indels: 262
DB: 10 Gaps: 32

US-09-914-168-2 (1-919) x US-60-453-134-135 (1-2349)
Oy 268 AspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGlu----- 285
    :::::::::::
    119 GATGTAATGACAGTTTACAGTATGACTTCTGTAAGCAGCATTTATAGTATCTGGGA 178
Oy 286 GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeu----- 302
    :::::::::::
    179 GGTATGCT-----GCCGAGCTTGAACTTACTTACCTGACGATTA 217
Oy 303 ---IleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsn 321
    :::::::::::
    218 CATTTAATATATCTTCCCGCGTAGTATGATGACAGTATGAAAATGCAATTAAGCA 277
Oy 322 AlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspValIle 341
    :::::::::::
    278 AAATCTGGAACACCGGTTAC-----GTTAGCGCAACGTTAAATTCAGTA----- 322
Oy 342 LeuProAsp-----AsnThrAlaAspValSerLeuIleTyrAspThrGly 356
    :::::::::::
    323 ---CCTGATTTGATGATGCAAAATTAACATTTAGCGATTAACCTTGTGTTGATGCTGGA 379
Oy 357 ThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGluLeu 376
    :::::::::::
    380 CGAGTTTAATCTGTTCCCAACTTCGCTTT-----GAGGAAATACCGCTT 424
Oy 377 ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal 396
    :::::::::::

```



Db 425 TCTGCTGAT-----AGCACTTTACGTCAGAAATGCCCAA 460  
QY 397 AsnMetGlyValAlaIleuValAsnLeuGlnAlaValAlaValAlaLeuSerAsnAspLeuIleAla 416  
Db 461 CAAGAAAGAACTTGGTATATTCACATTTAGT----- 493  
QY 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGln 436  
Db 494 -----GAGTTAGGAAAAATTCGC 511  
QY 437 AsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp 456  
Db 512 TTAGATCGTACAGCTTCTTCGAAACCGTGAACCGAATNGATCTCT----- 559  
QY 457 GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp 476  
Db 559 ----- 559  
QY 477 IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla 496  
Db 560 -----ATCAATGCTAGCAAT-----GATGAAGTGATGCTGA 592  
QY 497 AlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAsp 516  
Db 593 -----TATAAAGTCAAAAGACGTAAACACGGGTACTATCACTTT--- 631  
QY 517 AspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArg 536  
Db 632 ---GGTATTTGTTACGCTACAGAGAGTGGTATTTAGTTATACAGCACTGTT----- 679  
QY 537 AlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeu 556  
Db 680 -----AAACAAGATAATTTCTTAGACACAGCGCTCGAGTAAGTATA 721  
QY 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyrValPhe 576  
Db 722 GCTGGTACGAAA----- 733  
QY 577 ValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGly 596  
Db 734 -----AATGATTATGCT 745  
QY 597 ThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGly 616  
Db 746 ACAGAGTCTCAATTTGGTGTATACCGACCCCTAATTTTACTAAAGTGTGAAGCTTGGT 805  
QY 617 AlaGluLeu-----ArgLeuSerGluAspLys 626  
Db 806 GGAATATGTTTTCTTTGAACAACTACGATACCTTAATAAGTATACATCTCTTAATATAG 865  
QY 627 GlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAla 646  
Db 866 CGTACGACTATATGGAAGTATGTTACTTTAGGTTTCCCTGTAATGAATACCTCTAT 925  
QY 647 ThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666  
Db 926 TATGTAGCA-----TTAGGCCATACCTATATAT-----AAATATAGTAAC 964  
QY 667 ArgThrLeuGlnHisGluLeuSerArgSerIleIleGlnAsnGlyGlyTyrPAsnArg 685  
Db 965 TTTCCTCTAGAAATAT-----ACCCGTAAATTTATATATTCAA----- 1000  
QY 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuThrGlnAlaProProGluThrTyr 705  
Db 1001 -----TCAATGAATTTAAAGTAATGCAATTAACAAAT----- 1036  
QY 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuLeuAlaGly 725  
Db 1037 -----GACTTGTATTTTCTTTTGGTTGG----- 1060  
QY 726 ValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGln 745  
Db 1061 -----AACTATTAACAGCCTTATATAGAGCTATTTCCCA 1093

QY 746 ArgTyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsn 761  
Db 1094 ACTAAAGGGGTAAAGCAAGTCTTGGTGGACGAGTTACTTTCAGGTTCTGATTAACAA 1153  
QY 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGly 781  
Db 1154 TACTACAAACTAAGTCAGATGTACAGAGGTTTCTACCCATTA----- 1195  
QY 782 SerAsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSer 799  
Db 1196 ---GACGAGATCAACCTCTGGTGTATTCGAAAAACATCTCAGAGATATGCAAAATGCT 1252  
QY 800 AspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIle 819  
Db 1253 TTTCGAAACAAGCCTTTACCGCTTCTACCAACTATATACACGGGTGCACTGTTCAATTA 1312  
QY 820 ArgGlyTyrAlaHisAspSerLeuSerPro----- 829  
Db 1313 CGTGGTTTTCCTTATGTAGTACTATTTGACCTAACGCAATTTATGCCCTCAAAATGTAAT 1372  
QY 830 -----IleSerAspLysGlyTyrLeuThrGly 838  
Db 1373 AGTAATAGTAATAGTAATGCTACTTTTAATAGATAGTCTGAT-----GTGATTGCT 1426  
QY 839 GlyGlnValLeuAlaValAlaGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp 856  
Db 1427 GGTATGCAATCAACACTCGCAGCTGCGAAGCTTATTTACCACACCGCTTTGAGAGCAT 1486  
QY 857 -----LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869  
Db 1487 AAGGCCAAATACGTCGCAAGTCCCTTATTTGTTGATGCGGCAAGTGTGGAATACT 1546  
QY 870 -----AspLys 871  
Db 1547 AAGTGAATTCGATTAAGACTGGTGAATACAGTATTAATAAACTTAACCGCATTAAT 1606  
QY 872 GlyPheThrAsnAspThrLysIleGlyAlaGlyValAlaArgTyrPAlaSerProVal 891  
Db 1607 GGCAAATCAACCGGTATTCGCCGCTCTACAGGTGTGCGATTCCCAATGCAATCTCCATAT 1666  
QY 892 GlyGlnValArgValAspValAlaThrGlyValLys-----GluGluGlyAsnProIle 909  
Db 1667 GGGCCATTTGGTATTTCTCTTATGCCAAACCAATTAATAAATATGATGATGTCGAA 1726  
QY 910 LysLeuHisPhePheIleGlyThrProPhe 919  
Db 1727 CAGTTCCAAATTTAGTATTTGAGGTTCTTTC 1756

RESULT 31  
US-10-282-122A-27975  
; Sequence 27975, Application US/10282122A  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Heselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELIPIRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23





Oy	484	AlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla-----AlaLysAla	499
Db	1513	AAATTCACAGATTGGAAAAACGTTTGGCAACGGCAGCAACGGCAGCAACGC	1572
Oy	500	ArgHis-----	501
Db	1573	CATCATATGAAACTGTCGTACAAAGTAACATGATCCGAAATTGGCAATTCCTGCT	1633
Oy	502	-----LeuTyrAspMetProAspAspArgValLeuAla	512
Db	1633	CGCGTGCAGCGGATTCCTGTATCCACAAATTTGGACAGTGGAAAAAGATTAACCTTCTGAC	1692
Oy	513	ILeasnhIsaspaSpGlyValAsnaArgSerIleLeuGlyArgIleSerAspAlaValSer	532
Db	1693	ATGACGACATGA-----TTACATAAACCGTCATCGGCGAA---AATGAGCTGTGAG	1743
Oy	533	AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu	552
Db	1744	CGCTATCTATTAATGCAAT-----	1761
Oy	553	ArgThrAlaLeuAlaAsnaArgLysThrProAlaAspValTyrGlnSerLysValPro	572
Db	1762	CGTGGTACCCCTCGCGGATATTCGATTCGA-----AATGTCCTCGATGAT	1806
Oy	573	LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrrpGly	592
Db	1807	TCTTTTATGTTTTTG-----GGGCCACAGCGGGTAGCT--	1839
Oy	593	SerAspThrLysThrArgLysValThrLysPheGluHisAsnLeuLeuAsnArgAspGly	612
Db	1840	-----AAACCGAATTTGGCAAAAGCGTGGCCAAATTCCTATTTCACAGTGAAT--	1887
Oy	613	TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyrAla	632
Db	1888	---CATCGCATGCGCCCTAGACATGTGCGAATTATGAA-----	1926
Oy	633	ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln	652
Db	1927	-----AACACTACTGTTTCACAGCTGGTGGGGCTCCCTCGTGGATAT-----	1966
Oy	653	GluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGlu	672
Db	1969	-----GTGGATATGCAAGAGCGCGGTAT-----CTGACACAAAGCG	2004
Oy	673	IleSerArgSerIleGlnAsnGlyLysTrrpAsnArg-----	685
Db	2005	GTGCGCTGCTGCGCTTATTTAGATGATTTTACTGTGATGACATAGAAAAAGCTCATCTCGAT	2064
Oy	686	-----ThrTyrSerLeuArgTyrTrr-----ArgLeuAspLys	695
Db	2065	GTCTTCAACATCTGTACAGGATATTGGATGATGTCGTCGACGCGATGGACAGC	2119
RESULT 33			
US-10-282-122A-33378			
Sequence 33378, Application US/10282122A			
GENERAL INFORMATION:			
APPLICANT: Wang, Liangsu			
APPLICANT: Zamudio, Carlos			
APPLICANT: Malone, Cheryl			
APPLICANT: Haselbeck, Robert			
APPLICANT: Ohlsen, Karl			
APPLICANT: Zyskind, Judith			
APPLICANT: Wall, Daniel			
APPLICANT: Trewick, John			
APPLICANT: Carr, Grant			
APPLICANT: Yamamoto, Robert			
APPLICANT: Forsyth, R.			
APPLICANT: Xu, H.			
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms			
FILE REFERENCE: ELIURA.034A			
CURRENT APPLICATION NUMBER: US/10/282,122A			
CURRENT FILING DATE: 2003-02-20			
PRIORITY APPLICATION NUMBER: 60/191,078			

```

PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/230,335
PRIORITY FILING DATE: 2000-09-06
PRIORITY APPLICATION NUMBER: 60/230,347
PRIORITY FILING DATE: 2000-09-09
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/255,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/267,636
PRIORITY FILING DATE: 2001-02-09
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
Remaining PRIORITY Application data removed - See File Wrapper or PALM
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3378
LENGTH: 2763
TYPE: DNA
ORGANISM: Pseudomonas syringae
OS-10-282-122A-33378

```

Alignment Scores:	
Pred. No.:	0.0439
Score:	150.00
Percent Similarity:	32.65%
Best Local Similarity:	18.33%
Query Match:	3.17%
DB:	9
	Length: 2763
	Matches: 187
	Conservative: 146
	Mismatches: 355
	Indels: 332
	Gaps: 48

US-09-914-168-2 (1-919) x US-10-282-122A-33378 (1-2763)

[illegible]

Oy 184 LysAlaLysPheTyrGlnSerSerGlnSerGlnGluThrSerAlaIleGlySerSerHis 203  
 Db 571 CCCCC-----GTTGGCAAAAACCGCTGTGGC----- 600  
 Oy 204 GlnTyrThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSer 223  
 Db 601 CTGCTGGTGGGGGTGGGGGATCAAGAGCTTTACGACAACTGGAAATGGTGGCG 660  
 Oy 224 AlaMetAspLeuAsnGlySerIleProAlaGluArgGlnThrAlaLeuValAlaAlaArg 243  
 Db 661 ACCCTGGCGATCCCGGGGCC-----AAGACCTGGCCCGCAACCTTCAGAACACACCGC 714  
 Oy 244 AlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAsp 263  
 Db 715 GCGATGGCTTCTCTCTTACGATTTGGGACCAATCAG-----ATCGAC 759  
 Oy 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArg 283  
 Db 760 GTG-----CCGCTGGACATCGACCTGGAGACATTCGATTGC--- 795  
 Oy 284 GlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAsp---GluValProLeuLeu 302  
 Db 796 AGCGAGCCAGACCGCGAGCAAGCTGATGAACCTGACGCCGACCTGGAATTCAGAGCTGG 855  
 Oy 303 IleGlyAspValPheHisHisGlySerGluThrLysLysAsnLeuIleGluAsnAla 322  
 Db 856 ATCGAAGACCTGCAACGGGATGCCAAGCGGACGCGACAGAGCTGGTGTCCAGACGGC 915  
 Oy 323 SerAla-----GlnHisGlyTyrPheAspGly 331  
 Db 916 ACCGTGGCAGCCCAAGAACAGCAGTACGAGTATCTTGGACAAAGGCGAGTTCCGACGCC 975  
 Oy 332 ArgTyrLeuAspArg-----Ser 337  
 Db 976 ---TGGCTGAAGAAAGCTTCAGCGGACCGCTTTGCGCTTGACCCAGACAAAGCGC 1032  
 Oy 338 ValAsp-----ValIleLeuProAspAsnThr 346  
 Db 1033 ACCGACGCCACGCTGGCGCACTGGTGGGCTGTCTTGGCATCCAGCCCAAGACGCC 1092  
 Oy 347 AlaAspValSerLeuIleTyrAsp---ThrGlyThrGlnTyrArgPheAsp---GluVal 364  
 Db 1093 GCGTCACTCCCGCTACCCACTCTATATGAGTGGCCGACCACTGGACCTGGCGACACC 1152  
 Oy 365 ValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuPro 384  
 Db 1153 GTCTCAAGACGCTCAAGCG-----CTACTGGAAGACCCGAACAAGATCAAG 1200  
 Oy 385 ValLysArgGlu-----LeuLeuGluGlnLeuLeuThrValAsn 397  
 Db 1201 GTCGGTCAAGACGCCCAATTTGCCATCAACCTGCTGCCAATGCCCGCATTCGGCGTAT 1260  
 Oy 398 MetGlyGluAlaTyrAsnLeuGlnAlaValArg-----AlaLeuSer 411  
 Db 1261 CAGCGCGACGGCATCGAGCTGCGAGGTGGCTTTGACACCATTTCTGAATCTGATACGTA 1320  
 Oy 412 AsnAspLeuIleAlaThrArgTyrPheAsn-----MetValAsn 424  
 Db 1321 CTGGACTCCACCGCACTGCGACGACGCGGACAGCTGTGGGCCAAGTACTGACGCAAT 1380  
 Oy 425 ThrGluIleValPheProGlu-----ArgGluGlnIleGlnAsnAspGln 439  
 Db 1381 ACCGGCATCACTTCACAGACATCGCGCGGACAGGGTGCCACAGACAGTGCATTCGACGAG 1440  
 Oy 440 ValSerPheGlnGlnSerSerSerArgThrGluProAla----- 453  
 Db 1441 ATCGGCATCGACAGCGCGGTACTACGCGCGGAAGAGCGACCTGACCTGCGCGCTG 1500  
 Oy 454 ---GlnValAspGlnSer-----ThrLeuGluProValIleGluThr 466  
 Db 1501 CATGAAGTCTTCGAAGCCCGCTGGCGCGGATTCGACCTTCGACCGGCTGTCAACGAG 1560

Oy 467 ValGlu-----LeuThrAsp 471  
 Db 1561 ATCGAAGATGCCATTTGTCGGCTGCTGGCCCGCATCGAAGCTCAGGCGCGCTGCTCAT 1620  
 Oy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491  
 Db 1621 GCCAACCTGTGGCATCTCAGAGCGTTGAG-----CTGGCGGACAG 1662  
 Oy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeu-----TyrAspMetProAsp 507  
 Db 1663 ATGACCGCAGTCGAGCGGAGCGGCTTGGCATCGCCGGAAGAAATTCACCTCGGTTGC 1722  
 Oy 508 AspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyThrGly 527  
 Db 1723 CCCAAGCAGCTTGGGGTGATTTCTTACGAAAACTCGGATCGCATCTCTGACGAAC 1782  
 Oy 528 SerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluVal 547  
 Db 1783 GCCACCGGCGACCATCGACAGCCGAGGGGCGTGTGGCGGACCTCGCGAACAGAGAC--- 1839  
 Oy 548 IleAspLeuProGlu-----ArgThrAlaLeuAlaAsnArgLysThrPro 562  
 Db 1840 TTCCTCGTCCCAAGGTCTGATGACGATACCCCTGGATGAGCAAGCTCAAGACCTAT 1899  
 Oy 563 AlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysPro 582  
 Db 1900 ACCGAC-----CGCTGCCCGACAGATC-----AACCG 1929  
 Oy 583 ArgAspGlyGlnIleGlyLeuGlyTyrPglYserAspThrGly---ThrArgLeuValThr 601  
 Db 1930 CCGACCGGGCGGTATTCACACCTCTATCATGACGGCGGTAGCCGCTGACCGGGCTTGTCT 1989  
 Oy 602 LysPheGlnHisAsnLeuLeuAsnArgAspGlyTyrGlnAlaGlyValaGluLeuArgLeu 621  
 Db 1990 TCCAGCAGCCGCAATCTGCAGAAC-----ATCCCTATTGTGTACC 2028  
 Oy 622 SerGluAspLys-----LysGlyValLysLeuTyrAla 632  
 Db 2029 GCCGAAGCGCGCGCATCGCTCAGCGGTCTGTGGCGCCCAAGGTATCAACGCTGGCGG 2088  
 Oy 633 Thr-----LysProLeu 636  
 Db 2089 GCAGACTATTGCGAGATTGAACTGCGGATCATGGCGACCTGGCCAAAGACAAAGCGTTGG 2148  
 Oy 637 SerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGly 656  
 Db 2149 CTGCATGCGTTCGCAACACCTGGACGTCCACCGCGCACAGCGTGGCAAGTTTTCGCG 2208  
 Oy 657 HisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSer 676  
 Db 2209 GTCGAAGCTGAGACGTCACCAACCGCATGGCGCGAGCGCAAGCGATCAACTTCGCG 2268  
 Oy 677 IleIle-----Gln 679  
 Db 2269 CTGATCTACGGCATGACGCGCAATTTGGCTGGCCAAACAGATCGCGTGGACCGAAGCAG 2328  
 Oy 680 AsnGlyGlyTyrAsnArgThrTyrIleLeuArgTyr-----ArgLeuAsp 694  
 Db 2329 TCGAGCGCTATGTGATCGCTATTTTGGCCGCTACCCCGGGGTGCTGGCATATATGGAG 2388  
 Oy 695 LysLeuLysThrGlnAlaProProGluThrTyrPglAspLeuProValAspPheValAsn 714  
 Db 2389 CGACGCGCGACCGACCGCGCCGACGAG-----GGCTTCGTGAG 2427  
 Oy 715 GlyLysProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAsp 734  
 Db 2428 ACTATTTTGGCGGTGCGCTGTACTCGCGGACATC----- 2463  
 Oy 735 AsnLeuValAsnPro---MetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753  
 Db 2464 AACGCCAAGAACCTTCACTGCGCAAGCGCGCAACGCAATGCGCATCAACCGCGCATG 2523  
 Oy 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773

```
Db 2524 CAGGCGACGGCGCGC-----ATCATCAAAAAAGCATGTGGCGGTGAAT 2571
Oy 774 SerpHcglYAspAaAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyLeuGln 793
Db 2572 GCGTGGCGTGAACGAATCC-----GGCTTGAT 2598
Oy 794 AlaGlyTyrLleTyrPseAspAsnPhaHisValProTyrArgLeuArgPhePheAla 813
Db 2599 GCACGGGTCACTGTCACAGCTCAC----- 2622
Oy 814 GlyGlyAspGlnSerLleArgGlyTyrAlaHisAspSerLleuSerProLleSerAsp--- 832
Db 2623 -----GATGAACCTGGTACTGGAAGCTGGCGGAGACCTGGTTGATCAGATCAGCGAGCAG 2676
Oy 833 ---LysGlyTyrLeuThrGlyGlyGlyValLeuAlaValGlyThrAlaGluTyrAsnTyr 851
Db 2677 ATCGGTCCACACATGAGCGGAGCGGACAGACTGGCAGTG----- 2715
Oy 852 GluPheMetLysAspLeuArgLeuAlaValPheGlyAspLleGlyAsnAlaTyrAspLys 871
Db 2716 -----CCGCTGCTGCTGGAAGTGGCGGTGGCACACACTGGGAGCAG 2757

RESULT 34
US-10-263-929-30
: Sequence 30, Application US/10263929
: GENERAL INFORMATION:
: APPLICANT: Kim, Jaeseob
: APPLICANT: Galant, Ron
: TITLE OF INVENTION: Alzheimer's Disease Linked Genes
: FILE REFERENCE: USD-07417
: CURRENT APPLICATION NUMBER: US/10/263,929
: NUMBER OF SEQ ID NOS: 213
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 30
: LENGTH: 4802
: TYPE: DNA
: ORGANISM: Drosophila melanogaster
US-10-263-929-30

Alignment Scores:
Pred. No.: 0.119 Length: 4802
Score: 148.50 Matches: 190
Percent Similarity: 33.50% Conservative: 146
Best Local Similarity: 18.94% Mismatches: 376
Query Match: 3.14% Indels: 291
Caps: 48

US-09-914-168-2 (1-919) x US-10-263-929-30 (1-4802)
Oy 8 AlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLleuProLeuMetThrSerGln 27
Db 1867 GCAACAGACGCTACAAATCCTGCTGCCACACACAGACTTCTGACCATCCAGCAACAG 1926
Oy 28 ---AlaLeuAlaGlnGlnAsnAsnProAlaAsnLleLleAsnHisValProAlaHisAsp 46
Db 1927 TACAGATACCTACCAGCACAATCTGCTCCATG-----AAGTTAAGCGCTGGGC 1980
Oy 47 ThrAlaLleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnGlnLle 66
Db 1981 AGCACACTTATAGAAATCG-----CTGTTGATGGCTCGAGAGATT 2025
Oy 67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu----- 84
Db 2026 GATGCT-----TCTGTAGAGCGCTCAATCTGAAGCCCACTGTAACGATTAGTTATC 2079
Oy 85 -----AspValValAsnPheAspAspGlnSerProLleSerArgLleGlyGlu----- 100
Db 2080 AAACCCAAAGAGTAAAGTGAAGGTGAAGTCCCTTACTTATTTAGAGTCTCC 2139
Oy 101 -----GlnSerProProLeuGlyLeuAspMetSerValLleGluGluThrThrPro 117
```

```
Db 2140 AACACGCCCCAATCTCGACCCAAAGC-----GCAATCTCC 2175
Oy 118 LeuSerLeuGluGluLeuPheAla----- 125
Db 2176 AACAAAGAACCGAGCTGTTTAGCGACCATTCCAACTAGACCGGTACCAACAGCTGGA 2235
Oy 126 -----GlnGluSerThrGlnMetGly----- 132
Db 2236 AACTCGCTTGAGACACGAATGAGCGGAAGTCAAGATACAGCGCCCGTGAATCATGC 2295
Oy 133 IleAsnProAsn-----AspTyrLleProGluTyr-----Gln 143
Db 2296 CTGACTCTTAACAACTTTGAAGAGTACGCCAGCAACATTCAGACTGAGTGAATGATAG 2335
Oy 144 GlyGluGlnProAsnSer-----GluValValValProProThrLeuGluProGlu 160
Db 2356 GGATCGCCCATTAACAGACAGCTCAACAGAGCTTGCCAGCTTAACCTTTGATTAACATAC 2415
Oy 161 LysProGlyLeuLleLysArgLeuTyrAlaArgLeu-----PheAsnAsp 175
Db 2416 CGACCATCTCCACCGCTTCGCTGCTCAACCATCCAGAAATCCATTGTGAGAT 2475
Oy 176 GlyValAsnLysValProArgLeuLysValLysPheTyrGlnSerGlnSerGlyGlu 195
Db 2476 CAATCCACACCATTCACAGCTCGGAAAC-----TTCATTGGCACAGGCCAATGAG 2529
Oy 196 ThrSerAlaLleGlySerSerHisGlnLysThrGluProTyrAlaAsn---LleLysAla 214
Db 2530 AGTGCTCTTTCGAACAGATCTTAATGAGGACAGAGCTCGACGAACCAATCCCGCTTA 2589
Oy 215 AlaLeuGluAspLleThrGlnGlnSerAlaMetLysPheLysGlnGlySerLleProArgLeu 234
Db 2590 GCCATTGAGCGAGCTGCTGACAGGACAGACAGC-----GAATCA 2631
Oy 235 ArgGlnThrAlaLeuValAlaAlaArgValGlyTyrTyrAspLleAspLysLle 254
Db 2632 CATCCACACGACATTTGTC-----CTGCGACGCTGGCTTCTTACAAATTTCA----- 2679
Oy 255 IleArgAsnSerLleGlyGluValAspValLleLleHisAspLeuGlyGluProValTyr 274
Db 2680 -----TCGTAAAGACGACTACGCTCTATCTGGCGGAGACGATCTGTGGT 2730
Oy 275 IleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspArgLysAlaPheThrThr 294
Db 2731 CCGAATCTT-----ACGCTGGCTGGAGAGATACGCAAAAGTCTTCTTGGCAG 2781
Oy 295 ValAlaAspGluValProLeuLleLleGlyAspValPheHisGlnGlyTyrGluThr 314
Db 2782 GAATGATGATGCTGGCTGGCTGACCTGATGATGATGATGATGATGATGATGATGATGAT 2841
Oy 315 LysLysAsnLleLleGluAsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrLeu 334
Db 2841 ----- 2841
Oy 335 AspArgSerValAspValLleLeuProAsp-----AsnThrAlaAspValSerLle 352
Db 2842 -----ATCATTTATCCGACGACGACGACGACGACGACGACGACGACGACGACGAC 2886
Oy 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrLleAspProLys 372
Db 2887 CTAAACGCTGACGCCGACGAGTACTTGGACCAATA-----TGGCGTGTGACAAAGCC 2940
Oy 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGln 392
Db 2941 AAGCAGCAAGCATCAAGAGCCACACAGCTTG-----*----- 2973
Oy 393 LeuLeuThrValAsnMetGlyGluAlaLysValLysLeuGlnAlaValAlaGluLeuSerAsn 412
Db 2974 -----CTTGAGATGAGCTGGAGCGCAAGTTGGCGCGGCTG-----TGGGATAAAC 3021
Oy 413 AspLeuLleAlaThrArgTyr-----PheAsnMetValAsnThrGluLleValPhe--- 429
Db 3022 GAT-----ACGCGCTTATAGAAATACCGCCCGGAACGAGTACTGGGTCTTCCGC 3072
```

QY	430	-----	ProGluArgGluGlnIleGlnAsn	437
Db	3073	GTCAAGACACTTTTCTAAATAATGCGCTGGAGACAGCAGCAGCAAGACGCGATGGCCACT	3132	
QY	438	AspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGlu	457	
Db	3133	GACCCGMAAAAGCCAAATAATTCGACACGCTTGAGGCTCAGCAGCGGGGCAACGCAAGAAAG	3192	
QY	458	SerThrLeuGluProValIleGluThrValGluLeuThrAsp	471	
Db	3193	ATGAGCGCTCAATTCGCTGGCCCAAGGCAAGCAAAATTTCAAAATATGCAATGACGAGCCGTAACCTG	3252	
QY	472	-----	GlyIleLeuMetAspIleSerProIleGluPheSerAla	484
Db	3253	GACCCAAAAGCCTTGGTCCGAGGTGTGGCCAGCGGATTCAGGCCAATGATGATACCGCC	3312	
QY	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg	501	
Db	3313	GAGTTCCTTTTATATGACAAACACAATTTTTCAGAACTGGTGGCACTCCGATTTCTCC	3372	
QY	502	LeuIlyAspMetProAspAspArgValLeuAlaIleAsnHisAspArgValAsnArg	521	
Db	3373	ATGTTTGATTCACCGGGGTCAAGCTCCACAGATTACCAAGTCCACAGCTGTTTGGCTCAG	3432	
QY	522	SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAsp	541	
Db	3433	GAATGCGTAGCCACACAGGCGCCAAATAATGCACATATGAAGTCATCTTGTGTGGAG	3492	
QY	542	GluSerGluAsnGluValIleAspLeuProGluArgThrAla	558	
Db	3493	GACAAAGCTCCCGAG-----GACGAACCATGAGACTAGCGGGCGCTCTTCTTCGTCAT	3546	
QY	559	ArgLys-----	Thr	561
Db	3547	CGCAATCTTTTCAACGTGGAACCTCTTGTGGAAAGATGCGCTCGACGAGTTCACGT	3606	
QY	562	ProAlaAspVal-----	TyrGlnSerLysValProLeuTyrValPheValAlaSer	579
Db	3607	CAGTACGATTTTGAAACACCAACCATGCGCGCTCGCGATTTCTTCCTCCGCTCGAGGCC	3666	
QY	580	-----	AspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThr	595
Db	3667	AGCGTATGCTGATGTCACACTTTTGAAGACACATCGTCGATCGCACTGGACATGGACGATTTGTT	3726	
QY	596	GlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAla	615	
Db	3727	GCAGCTGTAAAGAAAGAAAGAAATTTGAATTCCTGCTACTAAACCTTCAGATTTGTTCTGT	3786	
QY	616	GlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyrAlaThr	633	
Db	3787	AAAGCCAAAGTG-----	GCTCCGATTAAAGTGGCTGCGACACGCGTT	3828
QY	634	-----	LysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr	650
Db	3829	CCCGTACCCAGGCTCATCGCTTATGAGATCGCGATATATGATCGCCGATTTGGGATTC	3888	
QY	651	GlnGlnGluValPheGlnHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu	670	
Db	3889	-----	TCAAGAAGCAGCGAGCTTTAAGCTCAGTTTGGTCA	3924
QY	671	HisGluIleSerArgSerIleIleGlnAsnGlyLysIleAsnArgThrTyrSerLeuArg	690	
Db	3925	-----	CAAAACATCATTTGGTACTGCCAAGCAGCATATATGAATGCAAAACCTTAA	3975
QY	691	TyrArgLeuAspLysLeuTyrThrGlnAlaProGluThrTrpGlnAspLeuProVal	710	
Db	3976	-----	GAATTTACTGGACATCCCTCGCGCTC	4002
QY	711	AspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys	730	
Db	4003	TCCATGATTTTTCGCCGCGCTCGGCCACGATTTG-----TCGCCCAAGTATGACAG	4056	

QY	731	ThTAlAlAAsrAsnLeuValAsnProMetAluGlyTyrAluGlnAlaTyr-----Ser	748
Db	4057	TTGGAGAGTTTAAACATGACGTCAAAGCAGCAAGCGGTTTTGAGAAATGATCATTTCCCAT	4116
QY	749	LeuGlnuValGlySerSerGlyLeuValSer-----	758
Db	4117	CTCGAAGTTGAGCTTAACAGCATGCGCTCTCTTAATGTGGAGGATGATGAGTGGCCGTCT	4170
QY	759	-----AspAlaAsnMetAlaIleAlaAlaGlyIleSerGlyValTyrSerPhe	775
Db	4177	ATACATCCCGCATACGCGTACGAAACTAGTT-----	4206
QY	776	GlyAspAsnAlaTyrGlySerAsnAlaGlnAlaHisGlnMetThrGlyIleGlnIleAlaGly	795
Db	4207	-----TCCAGACCTTTCTCCGAGTCTTTAACACAGAAATGTCTGCTTAACAGAC	4260
QY	796	Tyr-----IleTrpSerArgAsnProPheAsnHisValProTyrAluGlnuArgPhe	812
Db	4261	TATACCGTCGATCGTGTGTCG-----CTTTTATTCCCA	4293
QY	813	AlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp	832
Db	4294	CTATGGGCGCAT-----CAGCATGAACCTGGTGCACCTGGAAAG	4332
QY	833	LysGlyGlyTyr-----	835
Db	4333	AACCTCATTTACATGATGATGTGCGGCGCAATCTGATTCGATGCGCTTGAGAACAG	4392
QY	836	LeuThrGlyGlyGlnValLeuAlaValAlaGlyThrAlaGlyTyrAsnTyr---GluPheMet	854
Db	4393	CTGCTGGGCAAAAGATCTACTCTCGAAAAAAGTTAGACACCACAGCTACTGAGAGCATG	4455
QY	855	LysAspLeu 857	
Db	4453	TTTGATTTTA 4461	
RESULT 35			
US-10-267-502-24			
/ Sequence 24, Application US/10267502			
/ GENERAL INFORMATION:			
/ APPLICANT: Kim, Jaeseob			
/ APPLICANT: Galant, Ron			
/ TITLE OR INVENTION: Obesity Linked Genes			
/ FILE REFERENCE: LSD-07416			
/ CURRENT APPLICATION NUMBER: US/10/267,502			
/ CURRENT FILING DATE: 2003-01-27			
/ NUMBER OF SEQ ID NOS: 439			
/ SOFTWARE: PatentIn version 3.2			
/ SEQ ID NO 24			
/ LENGTH: 4802			
/ TYPE: DNA			
/ ORGANISM: Drosophila melanogaster			
US-10-267-502-24			
Alignment Scores:			
Pred. No.: 0.119 Length: 4802			
Score: 148.50 Matches: 190			
Percent Similarity: 33.50% Conservative: 146			
Best Local Similarity: 18.94% Mismatches: 376			
Query Match: 3.14% Indels: 291			
DB: 9 Gaps: 48			
US-09-914-168-2 (1-919) x US-10-267-502-24 (1-4802)			
QY	8	AlaAsnArgSerPheMetProValAlaLeuAlaIleAlaTyrLeuProMetThrSerGln	27
Db	1867	GCACACAAAGCTACAAATCCTGCTGCCCAACAAGCAAGCTTGGACCTGACCAACACAG	1926
QY	28	--AlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAsp	46
Db	1927	TACAAAGTAATACACAGCAACAATCTCGCTCCAAATG-----AAGTAAAGCGCTGGCG	1980
QY	47	ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnGlnIle	66

```

Db 1981 AGCACCTTATACMAAATCG-----CTGTTTGATGCTCTGACAGACTTT 2025
Oy 67 GlnAlaArgLeuAsnAlaIaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu----- 84
Db 2026 GATGCT-----TCTGAGAGGGCTTCAATCTGAGCCGCTTAACGATTGATTATC 2079
Oy 85 -----AspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGlu----- 100
Db 2080 AACCCCAAGTAAAAAGTGAAGGTGAAGAAATCCCTCTAGCTCTATTGGCAGTGTCCC 2139
Oy 101 -----GlnSerProProLeuGlyLeuAspMetSerValIleGluGluThrPro 117
Db 2140 AACACGCCCAATCTGACCCCAAGGC-----GCAACTGCC 2175
Oy 118 LeuSerLeuGluLeuPheAla----- 125
Db 2176 AACAAAGACCGAGGTGTTAGCGGAGCCATTCCAGTAGAGCGCTACCAACAGCTGGA 2235
Oy 126 -----GlnGluSerThrGluMetGly----- 132
Db 2236 AACTGCCCTGAGCAACGAATGAGGGAAGTACGATACGCGCCGCTGACTCATGG 2295
Oy 133 IleAsnProAsn-----AspTyrIleProGluTyr-----Gln 143
Db 2296 CTGATCTTAACAACCTTGAGAACGTACGCCAGCACAACTTCGACTGGAATGATGATCAG 2355
Oy 144 GlyGluGlnProAsnSer-----GluValValValProProThrLeuGluProGlu 160
Db 2356 GCATCGCCCATTAACACACACCTTCAGACGCTTGCCACAGTAACCTTTAGATACATAC 2415
Oy 161 LysProGlyLeuIleLysArgLeuTyrAlaArgLeu-----PheAsnAsp 175
Db 2416 GCACCAATCTCCACCCGCTCTCGCTGTCACCACTCCAGAAATTCATTGAGAGT 2475
Oy 176 GlyValAsnLysValProArgLeuLysAlaLysPheThrGlnSerSerGlnSerGlyGlu 195
Db 2476 CAATCCAGCACCATTCGACGCTCGGAAAC-----TTCACTTCGACGACGAGCCAAATGAG 2539
Oy 196 ThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsn-----IleLysAla 214
Db 2530 AGTGTCTTTCGAACACATCTAATGAGCGACGAGCTCTGCAGCAACCAATCCCGCTTA 2589
Oy 215 AlaLeuLysAspIleThrGlnGluSerAlaMetLysPheLysnGlnSerIleProArgLeu 234
Db 2590 GCCATTGAGCAGCTGCTGACAGGCGACGACGAC-----GAATCA 2631
Oy 235 ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle 254
Db 2632 CATCCAAACAGGCAATTCG-----CTCGACGCGCTGCTTACTATACAAATTTCCA----- 2679
Oy 255 IleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGluProValTyr 274
Db 2680 -----TCCGTAGACGACCTACGCTCTATCTGCGCGAGCGAGATCCTGTGTGTT 2730
Oy 275 IleAspTyrArgAlaValIleGluValArgGlyGluGluValAspLysAlaPheThr 294
Db 2731 CCGCAACTTT-----ACGGTGGGTGCGGAAGGATACCGCAATGCTTCTTCGGCAAG 2781
Oy 295 ValAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThr 314
Db 2782 GAATGATGTGGCTGCGCTGCAACCTCGATGACATTTCTCCACTTTCGAACAGAGATTT 2841
Oy 315 LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPhe 334
Db 2841 ----- 2841
Oy 335 AspArgSerValAspValIleLeuProAsp-----AsnThrAlaAspValSerLeuIle 352
Db 2842 -----ATCATTTATCCGACGACGAAACAGCCGCCCATTTGGTCAAGGT 2886
Oy 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLys 372
Db -----TACAAAGACGACGCTTTAAGCTCAAGTTTGGTCCA----- 3924

```

```

Db 2887 CTAAACCGTGACGCCGACGTCACCTTTTGACCACAGTA-----TGGCCGTTGACAAAGACC 2940
Oy 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGln 392
Db 2941 AAGCAGCAAGCATCAAGACCCACAAAGCTTG----- 2973
Oy 393 LeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsn 412
Db 2974 -----CTTGATGATGACGTGAGGAGGCAAGTTGGCGCGGTG-----TGCATAAAAAC 3021
Oy 413 AspLeuIleAlaThrArgTyr-----PheAsnMetValAsnThrGluIleValPhe--- 429
Db 3022 GAT-----ACGCGCTTTATGAATACCGCCGGAACCGGCTAGTGGCTCTTCCGC 3072
Oy 430 -----ProGluArgGluGlnIleGlnAsn 437
Db 3073 GTCAAGACATTTCTAATATATGCGCTGGAGACAGCAGCAAGAGAGAGATTGCCACT 3132
Oy 438 AspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGlu 457
Db 3133 GACCCGAAAAAGCCAAAAATTTGCACACGCTTGAGGCTCAGACGCGGCAACGCAAGAAAG 3192
Oy 458 SerThrLeuGluProValIleGluThrValGluLeuThrAsp----- 471
Db 3193 ATGACGCTCAATTCGTCGCCACAGCCGAGAAAGATTTCAGAAAGTACGACCGCTAAGTTC 3252
Oy 472 -----GlyIleLeuMetAspIleSerProIleGluPheSerAla 484
Db 3253 GACCCAAAGACCTTGTGCGCAGGTGTTCAGCGGATTCAGCGCAATGATGATACGCC 3312
Oy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg-----His 501
Db 3313 GAGTTCCTTTAATGAGCAAAACAAATTTTCCAACTGTGTGCACTCCGATTCTCC 3372
Oy 502 LeuTyrAspMetProAspArgValLeuAlaIleAsnHisAspArgLysValAsnArg 521
Db 3373 ATGTTTATCCACCGCGCTCAACGATTCACAGATTACCGTCCACAGCTTTTGCTCAG 3432
Oy 522 SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAsp 541
Db 3433 GAATGCTAGCAACGAGCGACCAAAATGACAGCTAATGAAGTCATCTTGTGTGAG 3492
Oy 542 GluSerGluAsnGluValIleAspLeuProGluArgThrAla-----LeuAlaAsn 558
Db 3493 GACAACTCTCCGAG-----GACGAACCCATGAGACTACGCGGCGCTTCTTCTGCTAT 3546
Oy 559 ArgLys-----Thr 561
Db 3547 CGCAAGTTTTCACAGCGAAGCCTTGTGGAAGAATGCTGCTCCGAGATTCAGT 3606
Oy 562 ProAlaAspVal---TyrGlnSerLysLysValProLeuTyrValPheValAlaSer--- 579
Db 3607 CAGTACGATTTTGAAACACCCATCGCCGCTGTGCGAATTTCTTCCCTCGCTGAGGCC 3666
Oy 580 -----AspLysProArgAspGlyGlnIleLeuGlyLeuGlyTyrPheLysSerAspThr 595
Db 3667 AGCCTGATGTGATGACATTTATGAGAGACATCGTGCATGCACTGGAAGCATTTGTT 3726
Oy 596 GlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAla 615
Db 3727 GCAGCTTAAGAAACAGAAATTTGAATAATGCTCTGCTCAATAAGCCTTCAAGTTTGTGT 3786
Oy 616 GlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyrAlaThr----- 633
Db 3787 AAGCCCAAGTG-----GCTCCGATTAAAGTTGGCTGCCACACGCTT 3828
Oy 634 -----LysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
Db 3829 CCGCTACCCAGGTCACATCGCTTATGAGATCGCGGATTAATGATTCGCCGATTTGGGATTC 3888
Oy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
Db 3889 -----TACAAAGACGACGCTTTAAGCTCAAGTTTGGTCCA----- 3924

```



```

QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyIleTrpAsnArgThrTyrSerLeuArg 690
DB 3925 -----CAAACTCATGTGTACTGCGCACACACTATAATACACTGCAAAACCTGAAA 3975
QY 691 TyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProVal 710
DB 3976 -----GAATTACTGAGACATCCCTGCCCGCTC 4002
QY 711 AspPheValAsnGlyLysProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLys 730
DB 4003 TCCATGATTTTGGCCGCCGCTCGGCCACCATTTG-----TCGCCAGTGAATGACG 4056
QY 731 ThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyr-----Ser 748
DB 4057 TTGGTGGAGTTTACATGCTCAAGGACACGAGGGTTTCGAGAAAGTATCATTTCCCAT 4116
QY 749 LeuGluValGlySerSerGlyLeuValSer----- 758
DB 4117 CTCGAGATTACACTTAACGACTGCTCTCATGTTAATGTGAGAGGTAGTAGTCCCGTGT 4176
QY 759 -----AspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe 775
DB 4177 ATCATCCCGATACGCGTACGAACACTGTT----- 4206
QY 776 GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGly 795
DB 4207 -----TCCAGCACCTTCTCCAGTCTCTTAAGCAGAGAAATGCTGCTTAAGAAAGAGAC 4260
QY 796 Tyr-----IleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePhe 812
DB 4261 TATACCGTGTCACTGTGGTGG-----CTTTTATTCGCA 4293
QY 813 AlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp 832
DB 4294 CTATGGGCGGAT-----CACGATGAACGTGTGACACTGGAAG 4332
QY 833 LysGlyTyr----- 835
DB 4333 AACTCTCATTTACATGCTGCTGTGGCGCAATCTCTATCCGAGTGGCTTGAGAACAG 4392
QY 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyr---GluPheMet 854
DB 4393 CGTGTGGGCAAGAACTACTCTCGAAAAAGTTAGCACACAGCTACCTGGAGCACATG 4452
QY 855 LysAspLeu 857
DB 4453 TTGGATTTA 4461

```

RESULT 36  
 US-10-172-502-1  
 : Sequence 1, Application US/10172502  
 : GENERAL INFORMATION:  
 : APPLICANT: FOSTER, Timothy et al.  
 : TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
 : FILE REFERENCE: P07263US01/BAS  
 : CURRENT APPLICATION NUMBER: US/10/172,502  
 : PRIOR FILING DATE: 2002-06-17  
 : PRIOR APPLICATION NUMBER: US 60/298,098  
 : NUMBER OF SEQ ID NOS: 29  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 1  
 : LENGTH: 6609  
 : TYPE: DNA  
 : ORGANISM: Staphylococcus epidermidis  
 US-10-172-502-1

Alignment Scores:  
 Pred. No.: 0.288 Length: 6609  
 Score: 146.00 Matches: 185  
 Percent Similarity: 34.148 Conservative: 138  
 Best Local Similarity: 19.568 Mismatches: 365

```

Query Match: 3.09% Indels: 258
DB: 9 Gaps: 40
US-09-914-168-2 (1-919) x US-10-172-502-1 (1-6609)
QY 27 GlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisasp 46
DB 4201 CAAGCATTTAGCTAAAGAAAGAAAGAAAGACACTT----- 4233
QY 47 ThrAlaIleAsnGlnAlaValAspAlaGlyAsnProProValLeuLeuThrProGluGlnIle 66
DB 4234 GCAGCTATTGACCAACT----- 4257
QY 67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
DB 4258 AATAGTCAGCTGATCATCAACGGCCACCAAAATGCT-----GTATCACGCTATAAAAT 4308
QY 87 ValAsnPheAspAspGln---SerProIleSerArg-----IleGlyGlnGlnSerPro 103
DB 4309 ATTCAACCTGAAACAAATAAACCAGCAGCAGCTGAAAAAATCATCAATAAAGCGAAT 4368
QY 104 ProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeu----- 4428
DB 4369 GAATTACGTGGCCAAATTATATCAAGATTAAAGCGACAGCAGACAGAACAGACGCGG 4428
QY 119 -----SerLeuGluGluLeuPheAlaGlnGlnSerThrGluMetGlyIleAsnProAsn 136
DB 4429 TTAGATTAATCATCATGATTAGTGTCTAAAGCATGACAAATATC-----ACGAAT 4479
QY 137 AspTyrIleProGluTyrGlnGlnGluGlnProAsnSerGluValValProProThr 156
DB 4480 GATAGACAATAATCGCAAGTATATGACTCACACAAACGCTTGACGACATTGCATT 4539
QY 157 LeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGly 176
DB 4540 GTGACCGCTGACCAT-----ATTGTTAGCAGCTGTAGA-----GATGCA 4581
QY 177 ValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThr 196
DB 4582 GTTAAAGCAA-----CAATATGAGCTAAAGACGAATGAGCAACGCAACGATCGG 4635
QY 197 SerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeu 216
DB 4636 ACTGATGAAGAAACAAACAGTCTTTAAATCAATTAGCAATATGAAAAAGCTGATT 4695
QY 217 GluAspIleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGln 236
DB 4696 CAAACATTATATCAAGCAATACG-----AATATGATGTGAACGTGTGAATCA 4746
QY 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg 256
DB 4747 AATGATATTGCTACGTTAAAGCGCTA----- 4773
QY 257 AsnSerIleGlyLysValAspValIleIleHisAspLeuGlyGluPro---ValTyrIle 275
DB 4774 ----- 4791
QY 276 AspTyrArgAlaValGluValArgGlyGluGlyAlaAspLys----- 290
DB 4792 AAACCTGAGCTCAAGCAACCAATAAAAGCGACGATACCAACATAGATCTATAAA 4851
QY 291 -----AlaPheThrValAlaAspGluValProLeuLeuIleGlyAspVal 306
DB 4852 GATACACCACTGCTAGCAGACATGATTAATGATGAAGCAACCAATAAAGCGACACA 4911
QY 307 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlnAsnAlaSerAlaGluHis 326
DB 4912 CTTAAACAAGGTCAACAAAGT-----ATGACAAATTCGACCAAGATGCA 4956
QY 327 GlyTyrPheAspGlyArg---TrpLeuAspArgSerValAlaIleLeuPro----- 343
DB 4957 GCTGTCAATGATGTGAAGAAACCAACGATTAAAGGCAATGCAATTAATTAACCGAAAGTT 5016

```

QY 344 -----AspansThralaAspvalSerleuIleTyraSphrGly 356  
Db 5017 AGACCGAAACGTGCAGCGTTGGATTAAC----- 5043  
QY 357 ThrGlnTyrArgPheaspGluValAlaPhepethrIleAspProLysThrAsnGlnLeu 376  
Db 5044 -----ATTGATGAGAAAGTAATAATAATCAACTGCAGTAATCGAATACGGTA 5091  
QY 377 ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal 396  
Db 5092 GATACACCGACGATGACGAAGAAATGCTGATTCGCGTTAAATAAAT----- 5142  
QY 397 AsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaGalaLeuSerAsnAspLeuIleAla 416  
Db 5143 -----GTTAATGCAATTAAATAAATGATTCACAA 5172  
QY 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlnAlaArgGlnIleGln 436  
Db 5173 -----ACAAACGCAATCCAGAAAGTGATCAAACTGAGCGGTGATTAACAC 5220  
QY 437 AsnAspGlnVal-----SerPheGluGlnSerSerSer 447  
Db 5221 AACATCAAAAGCATTTTACCTAAACTTCAAGTTAAACGACGCGGTCAATCTGTACG 5280  
QY 448 SerArgThrGluPro-----AlaGlnValAspGluSerThrLeuGluProValIleGlu 465  
Db 5281 GCAAAAGCTGAAGCTCAAAATGCACTTATGTATCAAGTGAATTTATCCGAGAAAGAA 5340  
QY 466 ThrValGluLeuThrAspGlyIleuMetAspLysSerProIleGluPheSerAlaSer 485  
Db 5341 -----AGATTAGCTCTTAA 5355  
QY 486 AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet 505  
Db 5356 CATTTAGTAGAACAAGCACTTAATCAACCTATGTATCAAGCAATCAACGCA----- 5406  
QY 506 ProAspAspArgValIleuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly 525  
Db 5407 -----GATTAGACTGCCCAACTTATCAAAATAGTATCGATGCTCAAAATTTATTTCA 5460  
QY 526 ArgIleSerAspAlaValSerAlaValAlaArgAlaIle-----LeuProAspGluSer 543  
Db 5461 AAAATTTAAACCGACGACAAAGTTAAACGACGATTAACAACAAATTCAAATATCGCT 5520  
QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563  
Db 5521 ACAATTAATAATTAATTAATAAGCAATATACGAACGACGATGAGAAACAAATGCT 5580  
QY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583  
Db 5581 GCATTAAGTACAAAGTTGAAAGAGTAAAT-----AAAGCTTAA 5619  
QY 584 AspGlyGlnIleGlyLeuGlyTyrPGLysSerAspThrGlyThrArgLeuValThrLysPhe 603  
Db 5620 CAAACAAATTCCTGCT-----GCAGTACTAAATGCT 5649  
QY 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgSerGlu 623  
Db 5650 GATGTCGATATTTATTCATGATGAGGAAAC-----GAAATTCGGAATCGAA 5700  
QY 624 AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGln 643  
Db 5701 CCTGTTATTAATAAAAGCAACGACCGCGACAAACATTTAACACATTTATTCACGATTAAG 5760  
QY 644 -----LeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGly 656  
Db 5761 AAACAAACAAATTGAACGAATGTTCAAGCAACGTAAGAAAGAAATAGATTTTGA 5820  
QY 657 HisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSer 676  
Db 5821 CAGTTACAAACATTTATGCACTGCTATTGACAAATTAATCAAGATCGTAGCAATGCA 5880  
QY 677 IleIleGlnAsnGlyLysTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeu 696

Db 5881 -----CAAGTGTATTAACCA 5895  
QY 697 LysThrGlnAlaProGluThrTyrPGLAspLeuProValAspPheValAsnGlyLys 716  
Db 5896 GCACAA-----TTAAATCTCAAAACAATACATGATTTAGACGTACATCCATTAATAA---AAG 5949  
QY 717 ProSerGlnGluAlaLeuLeuAlaGlyValAlaValaHisLysThrValAlaAspAsnLeu 736  
Db 5950 CCAGATGCTGAA-----AAACGATTAATGATGATCTT 5982  
QY 737 -----ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753  
Db 5983 GCACGTGTTACACATTTAGTCGCAAAATATGCAAAA----- 6018  
QY 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773  
Db 6019 -----GTAAGTGAATGTAATTAAGCTGATGCTTAAAGCTATTAACGTCAATTA 6069  
QY 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIleGln 793  
Db 6070 TTACAAATGATGAAGATTTAAAAACACACGCACTAATGCT----- 6111  
QY 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813  
Db 6112 -----GATGTTGATGCAGTTTAAACGATTTAAATGTTGCAATTA 6150  
QY 814 GlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833  
Db 6151 GGCAGTATAGAACAGATTAATTAAGTAAAGAAAGAAAGAAAGTAAAGTAAAGTAAAGTAA 6210  
QY 834 GlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPhe 853  
Db 6211 GCTCAACAAACATATGCGAAATTCGAAAGCGATCCGACACAGAA---CAATTAGCTTAA 6267  
QY 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe 873  
Db 6268 GTAAAGCATTAATTAATGATCAATATGTTGACAGATGCAATGAATGATGAAGATGCG 6327  
QY 874 Thr---AsnAspThrLys-----IleGly 880  
Db 6328 ACATTAATGATGACATCAAAAGATACGCACTCTATTGATGAAATTTATAGCAATTTAA 6387  
QY 881 AlaGlyValGlyValaGlyTyrPalaSerPro---ValGlyGlnValaArgValaAspValaAla 899  
Db 6388 TTACCTGCTGAGATGAATAAAGCGTCAACAAAGTGGGCAACCTGCTCCAAAGTTTGT 6447  
QY 900 ThrGlyValLysGluGlu 905  
Db 6448 ACGCCTATTAAAAAGAA 6465

RESULT 37  
US-10-320-800-39  
Sequence 39, Application US/10320800  
GENERAL INFORMATION:  
APPLICANT: ROBINSON, ANDREW  
APPLICANT: GORRINGE, ANDREW  
APPLICANT: HODSON, MICHAEL  
APPLICANT: REDDIN, KAREN  
TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE  
FILE REFERENCE: 1581.0790001  
CURRENT APPLICATION NUMBER: US/10/320, 800  
CURRENT FILING DATE: 2002-12-17  
PRIOR APPLICATION NUMBER: PCT/GB99/03626  
PRIOR FILING DATE: 1999-11-02  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 39  
LENGTH: 7542  
TYPE: DNA  
ORGANISM: Neisseria meningitidis  
FEATURE:  
NAME/KEY: CDS



```

Oy 585 yGlnlledlyLeucllylrprclserAsprThrlgylTrhArgleValThlLysPheGlnHl 605
Db 2499 CGArTATGCGCGTTCGGCGGCCCATTAAGTATACCAATGGCCACAGCCAGCCAAAGGCCAAAG 25588
Oy 605 sAsnleuLleAsnArgspclY-----TyrGlnAlaGlyAlaGluLeuArgLeuSerGcl 623
Db 2559 G-----ATACACAAATGCCCGCCCAACCATTCAGACCTCGCAGCCAAAAATCCGTTTGCGT 26122
Oy 623 uAspLyLeuGlyValLysLeuYrAlaThlLysProLeuSerHlIsProLeuAsnArgcl 643
Db 2613 AGAAAG-----CTGCACAAATACCAGAGCA 26399
Oy 643 nLeuArgAlaThrlleu-----GlyYrGlnGln-----GluValPheGcl 656
Db 2640 TTTCGAAACACGATTCGTAGAAACAGGCGCGCATATTGTTGATTACGAAGCATTTGG 26999
Oy 656 yHlSerThrAsnclYrheAspLeuSerThrlArgThrlLeuGlnHlSglLlIleSerArgse 676
Db 2700 AGCA-----CAGCAATTATTGGCAGAGCCAGCAACATGAATTAAGCTGCTC 27479
Oy 676 rlllellecln-----AsnGlyGlyTr 683
Db 2748 TGCTATACGATGATCATCACACCATTAAGCACCCTGATGAGAGGCGCGCATGAATAATG 28079
Oy 683 pAsnArgThrlYrIserleuArgYrArgleuAspLysleu-----LysThrlGln----- 699
Db 2808 GCAT-----AAATACATTAATGAAAGCCACCCAAACCAACCAAGATTAC 28529
Oy 700 -----AlaPro-ProGlnThrlrGlnAspLeuRoValAspPheValAsnGlyLys- 716
Db 2853 CCAACATCGCGCACCAAAATCATTTACAGTATAGTTTAACCATTTGATGTAAGAAGCT 29129
Oy 717 -----ProSerGlnAlaLeuLeuAlaGlyValAlaValHlSylsThrValAla 734
Db 2913 ATTTAAATACCATGATGCCAAATCATCTGTGTGGCAATTCATTTGACAAAGAAAGAA 29729
Oy 734 sPAsnLeuValAsnProMetArgGlyYrArgGlnArgYrSerLeuGluValGlySers 754
Db 2973 CGGTTTCATACAGCAAACTTTGGCGCAAAAGAAAGATTTCATGTAATATGCAAAAT 30329
Oy 754 erGlyLeuValSerAspAlaAsn-----M 762
Db 3033 ACACAGCTATTGGCGTACGAAACATAAAGACGACGACTCAACGGCATACGCAACAAA 30929
Oy 762 eAlaIleAlaArgAlaGlyLlIleSerGlyValYrSerPheGlyAspAsnAlaYrGlyS 782
Db 3093 TTACACTTTGCCGGAGGAA-ATACACAGCAACATTTCTAGTGGGTTCATTTGCTATGAT 31519
Oy 782 eAsnArg-----AlaHlSglN-----MetHrG 790
Db 3152 CGCATCGCAAGCATTTAAGCATTCATCGCGCCACCAAGCATAGTTCGCCCAAAACA 32119
Oy 790 lYcIlyllleGlnAlaGlyYrlllrPserAspAsnPheAsnHlSValProYrArgleuA 810
Db 3212 ACGGTATTTCGATCCCTATACG---TCCAAATCTTTTACCCCATTACCACACAGCACT 32689
Oy 810 rGpPheAlaGlyLysArgSlnSerLleArgGlyYrAlaHlAspSerLeuSerProI 830
Db 3269 TATATATT-----ATCAATCTCTG 32869
Oy 830 leSerAspLysGlyYrLeu 836
Db 3287 TC---AATAAGGCTATCTT 3303

RESULT 38
US-10-122A-29937
: Sequence 29937, Application US/10282122A
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Melone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlisen, Karl

```

```

APPLICANT: Zyskind, Daniel
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/220,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 29937
LENGTH: 7545
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-282-122A-29937

```

Alignment Scores:			
Pred. No.:	0.376	Length:	7545
Score:	145.50	Matches:	187
Percent Similarity:	34.818	Conservative:	150
Best Local Similarity:	19.328	Mismatches:	336
Query Match:	3.08%	Indels:	297
DB:	9	Gaps:	48

  

US-09-914-168-2 (1-919) x US-10-282-122A-29937 (1-7545)			
OY	19	AlAtyrLeuProLeuMeThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIle	38
		:       :	
Db	841	GCACATTCGGCTATTCACATTAATGCTGCGCCAAATACCTCAACAAATACAGCCACMAC	900
OY	39	IleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaIleValaGlyAsnPro---	57
		:       :	
Db	901	GGCACACATATCCCTTATTT-----GGCATTTGATACACGGCAAAATTAGAGCTTGTAT	954
OY	58	-----ProValLeuThrProGlnGlnIleGlnAlaArgLeuAsnAlaIa	73
		:       :	
Db	955	GCCAAACAAATACACCTTGATGACAGTACAGCCGACAGCAAGCAGCATTCGT---AATCAAGCG	10111
OY	74	GlyLeuAsnAlaIysProGlnInsertGlnAlaLeuAspValValAsnPhaSpasGlnInsert	93
		:       :	
Db	1012	CAGTTGTTGCTCTTCCTCCGCTAATGCGGAGTATGATGA-----AATGCC	1056
OY	94	ProIleSerArgIleGlyGlnGlnInsert-ProproLeuGlyLeuAspMetSerValIleG	113
		:       :	
Db	1057	CGTTTACATCATAGTGGCAGCAGATGCGTGGCCGCAATG---CCAAAGATACGATATATACA	1113
OY	113	uGluThrThrProLeuSerLeuGlnGlnAlaLeuPheAlaGlnGlnInsertThrGluMetGlyI	133
		:       :	
Db	1114	GCGGACACACAA--GTCAATATTCGC-----AGTCAAGGCGCT	1148

QY 133 easnProbaAspTyrIleProGluTyrGlnGluGlnProAsnSer----- 149  
 Db 1149 TCAAAACACACGGGTACGGCGGTATCCGACACAGGCACTCAAAATTCACAGTCAGTCATCA 1208  
 QY 150 -----GluValValVal-----ProThrLeuGlu 158  
 Db 1209 AACACGTGCACATATATGTCCTCCAGCGCAAAATATGATTCACAAATTCGGCGAGCTGAA 1268  
 QY 158 uProGluTyrProGluLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAs 178  
 Db 1269 AAATGAACATCAGCAGCATTGAAAGCCGCTGTTGGCTATGATATACGACACTTAA 1328  
 QY 178 nLysValProArgLeu-----LysAlaLysPheTyrGlnSerSerGln----- 192  
 Db 1329 TAAATCAGGCAAACTCTCTCAAAACAGTTCACAAAACCTCCATATGATGCAACAGCAA 1388  
 QY 193 -----SerGlyGluThrSerAlaIleGlySe 201  
 Db 1389 AATGGATAACCGTGGCCGATGGGTTTACAGATACCGCACCAACCGCGCTCAATGGTTC 1448  
 QY 201 rSerHisGlnLysThrGluProTyr----- 209  
 Db 1449 AAGCAATCAAAACCGCAATAGTTACATGCAATCTTCCATTCACATCCACACCAAC 1508  
 QY 210 -----AlaAsnIleLysAla----- 214  
 Db 1509 AACGGCAACAGGTACGGGTACTGCACACCGTTTCTATATCAACATACATACGCGCTACCTT 1568  
 QY 215 -AlaLeuGluAspIleThrGlnGluSerAlaMetAspLysAsnGlySerIleProArgLe 234  
 Db 1569 TCCTGTATGGACAAATTCGACATCATGTCACGTGATTAATTCAGCGCATATTTATGCCAA 1628  
 QY 234 uArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle 254  
 Db 1629 TGGTCAACAGATGTTAGTGGCGCAACAGGTTTA----- 1662  
 QY 254 eIleArgAsnSerIleGlyLysValAlaIleIleHisAspLeuGlyGluProValTyr 274  
 Db 1663 -----AATTAATGCAGCAATAATAC-----ATTCAATCAGTTA----- 1695  
 QY 274 rIleAspTyrArgAlaValAlaGluValArgGlyGluGlyAlaAspAspLysAlaPheThrTh 294  
 Db 1696 -----AATGCCAAAGGTTGCGGTTGCGCAATGCACAAATGGAACAAAT 1736  
 QY 294 rValAlaAspGluValProLeuLeuIleGlyAspValPheHis-----HisGlyLysTyrGlu 313  
 Db 1737 TATCACTGATGGGCTCCACATTCACCGCGACCTGAAATATCAAAATGGCAACATCAC 1796  
 QY 313 uThrLysLysAsnLeu-----IleGluAsnAlaSerAlaGluHisGlu 327  
 Db 1797 AACACGCCCAAGATTAGATTGAACCGCATCACTGATTAACGCT-----CATGG 1847  
 QY 327 y---TyrPheAspGlyArgTyrLeuAspArgSerValAspValIleLeuProAspAsnTh 346  
 Db 1848 CAAGTATTATTATCAGCAAAATACCGGATTTTCCAGCGAGCGCTGTAACATCAAAA 1907  
 QY 346 rAlaAspValSer-----LeuIleTyrAspThrGlyThrGluTyrArgPheAs 362  
 Db 1908 TGGCGAAATAAGCAGCAATCAACACTGATTTATTCAGATGTCACAAATCTAAC----- 1962  
 QY 362 pGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382  
 Db 1963 -----GCGTCTATGATTAATACGAATGCG----- 1986  
 QY 382 sLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr 402  
 Db 1987 -----ACGATCAATTCAGCGCGCTGATGT 2009  
 QY 402 rAsnLeuGlnAlaValArgAlaLeuSerAsnAsp-----LeuIleAlaThrArgTyrPhe 420  
 Db 2010 TGGTATTCAGGCA---AAATCGTTATCCAAACAGCGCACACTTGGCGCTGATTAATAACT 2066

QY 420 eAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGluVal 440  
 Db 2067 GGAATATGGGTTACAGATGATTTTATGATAGAACCAATATCGCGGGCAATAAAT 2126  
 QY 440 lSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspLysSerThrLe 460  
 Db 2127 GTGCGTCACTACACGAGCGCTGAAAATTC----- 2160  
 QY 460 uGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLysSerProIle 480  
 Db 2161 -----CATACTTGGACACGAGAAAACCATTCGGATTAAGCAATATACCT 2207  
 QY 480 eGluPheSerAlaSerAsnLeuIleGlnAsp-----LysLeu 492  
 Db 2208 TGATATGCAGCAGCAAGGCAACATTCAATCCGGGTACGACAGACATTGGCAGCAGCA 2267  
 QY 492 uAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAl 512  
 Db 2268 CAATTTAAC-----AATAGAGCTTGATTGACGCGACAAACCAAAATCCAAAC 2318  
 QY 512 aIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSe 532  
 Db 2319 C-----GGCAATGATATATATCGGTACAGATGCGATTTATGGCACAATAT 2366  
 QY 532 rAlaValAlaArgAlaIleLeuProAspGluSerGluAsn-----GluValIleAs 549  
 Db 2367 CGCTATTTGGCGTACCGCGGTATACAAATCAAGATGAAGAACGATACAGTCCCGCATTCG 2426  
 QY 549 pLeuProGluArgThrAlaLeuAla-----AsnArgLysThrProAlaAspVal 565  
 Db 2427 GGCACGTGAAACCTGAATTTAGCATCGACATTAACCAACCGTGAACATGCTGAT 2486  
 QY 565 lTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlu 585  
 Db 2487 TTRACACCGT-----AA 2498  
 QY 585 yGlnIleGlyLeuGlyTyrPGLysAspThrGlyThrArgLeuValThrLysPheGluHis 605  
 Db 2499 CGATATGGCGGTGGCGGCGCATTAATGATACCAATGGCCAAAGCCACAAAGCCCAAG 2558  
 QY 605 sAsnLeuIleAsnArgAspGly-----TyrGlnAlaGlyAlaGluLeuArgLeuSerGlu 623  
 Db 2559 G-----ATACACAAATCGCGCGCAACCATTTGAGCTGCAGGCAAAATGGCTTACGTGT 2612  
 QY 623 uAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlu 643  
 Db 2613 AGAAAG-----CTGCACAAATACCAATGACGA 2639  
 QY 643 nLeuArgAlaThrLeu-----GlyTyrGlnGln-----GluValPheGlu 656  
 Db 2640 TTGGAAGCGCGCTGTAGAAACAGGCGCGCATATTTGTTGATTTACGAACATTTGG 2699  
 QY 656 yHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSe 676  
 Db 2700 ACCA-----CAGCAATTTATGCGAAGGACACGACACATGAATTAGCTGTGTC 2747  
 QY 676 rIleIleGln-----AsnGlyLysTr 683  
 Db 2748 TGTCTATACGATGATCAGACCATTCAGCACCCCTGATGAGAGCGGCGCATGAATAATG 2807  
 QY 683 pAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeu-----LysThrGln----- 699  
 Db 2808 GCAT-----AAATACGATTTATGAAGAGTCACCCCAAAAACCCCAAGTTAC 2852  
 QY 700 -----AlaPro-ProGluThrTyrPGLAspLeuProValAspPheValAsnGlyLys- 716  
 Db 2853 CCAACTGGCGCGCAAAATCATTTTCAGGTAATGATTAACCATGATGTAAGAAAGT 2912  
 QY 717 -----ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAla 734  
 Db 2913 ATTTAATACGATGACCAATCATTCGTGTCGCAATCTCATTTGTACMAACGAAAAAGA 2972  
 QY 734 sPAsnLeuValAsnPrometArgGlyTyrArgGlnArgTyrSerLeuGluValGlySers 754

```

Db 2973 CGGTTTGCATACGACGAACCTTTGGCGAAAGAAAGTATTCAGTGAATAATGGCAATT 3032
Oy 754 eRgLYleuValSerAspAlaAsn-----M 762
Db 3033 ACACAGCATTTGGCGTGAAGAAACATTAAGACGACATCAACGGGACATAGCGAACAAA 3092
Oy 762 eAlaIleAlaIArgAlaGlyIleSerGlyValTYrSerPheGlyAspAsnAlaTYrGlys 782
Db 3093 TTACACTTTCCCGGAGAA-ATCACACGCAACATTTCCATTCGTTGCTATGAT 3151
Oy 782 eTAsnArg-----AlaHisGln-----MetHisG 790
Db 3152 CGCATCGCAAGCATTAAGCCATCATGCGCCGACGCAAGGACATGATGCCGCAAGCA 3211
Oy 790 LYgLYleGlnAlaGlyTYrIleTYrSerAspAsnPheAsnHisValProTYrAlaGlu 810
Db 3212 ACGGATTTCCGTCACCCATAGC---TCCATTTCTTTTACCCCATTTACCCAGACAGCT 3268
Oy 810 rGpHePheAlaGlyGlyAspGlnSerIleArgGlyTYrAlaHisAspSerLeuSerProI 830
Db 3269 TATACATT-----ATCATCTCTG 3286
Oy 830 IeSerAspLYsGlyTYrLeu 836
Db 3287 TC---AATTAAGGCTATCTT 3303

RESULT 39
US-10-018-470-1/c
: Sequence 1, Application US/10018470
: GENERAL INFORMATION:
: APPLICANT: Frazer, Claire M.
: APPLICANT: Hickey, Erin
: APPLICANT: Peterson, Jeremy
: APPLICANT: Tetteilin, Herve
: APPLICANT: Venter, J. Craig
: APPLICANT: Masignani, Vega
: APPLICANT: Galeotti, Cesira
: APPLICANT: Mora, Manrosa
: APPLICANT: Ratti, Giulio
: APPLICANT: Scarselli, Maria
: APPLICANT: Scarlato, Vincenzo
: APPLICANT: Rappuoli, Rino
: APPLICANT: Pizzi, Mariagrazia
: APPLICANT: Grandi, Guido
: TITLE OF INVENTION: Neisseria Genomic Research
: FILE REFERENCE: CHIR-0319
: CURRENT APPLICATION NUMBER: US/10/018, 470
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/132, 068
: PRIOR FILING DATE: 1999-04-30
: PRIOR APPLICATION NUMBER: PCT/US99/25373
: PRIOR FILING DATE: 1999-10-08
: PRIOR APPLICATION NUMBER: GB-0004695.3
: PRIOR FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: PCT/US/05928
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 2272325
: TYPE: DNA
: ORGANISM: Neisseria meningitidis
US-10-018-470-1

Alignment Scores:
Pred. No.: 609
Score: 145.50
Percent Similarity: 33.48%
Best Local Similarity: 20.49%
Query Match: 3.08%
DB: 9
Caps: 40
Length: 2272325
Matches: 191
Conservative: 121
Mismatch: 342
Indels: 278

```

```

US-09-914-168-2 (1-919) x US-10-018-470-1 (1-2272325)
Oy 116 ThrProLeuSerLeuGluIleuPheAlaGlnIleuSerThyGluMetGlyIleAsnPro 135
Db 180761 TCGCCTTTGGCACTTGGCGACTTCCACATCCAAAGACATCCGGGTGGAAGGCTTGACGGCT 180702
Oy 136 AsnAsp-----TYrIleProGluTYrGlnGlyGluIleProAsnSer 149
Db 180701 ACCGAGCGGAGTACCGTATTCATCACTGCGCCGCAAGTCCGCGACACCTACAC--- 180645
Oy 150 GluValValValProProThrLeuIleuProGluIleuProGluIleuIleuArgLeuTYr 169
Db 180644 -----GACACACAGGCACTGCCATCATCAAAAGCCTGTAC 180609
Oy 170 AlaArgLeuPheAsnAspGlyValAsnLYsValPro-----ArgLeuLYsAlaLYs 186
Db 180608 GCCACCGGTTCTTTGACGAGCTACGCGTCCGAACTGGCGAGGCGGACCTCTCGTGACC 180549
Oy 187 PheTYrGlnSerSerGlnSerGlyIleuThrSerAlaIleGlySerSerHisGlnLYsThr 206
Db 180548 GTTATCGAAAGCCGACCATGCGTCCGCTCAACATCACCGGCGCAAAATGCTGCAAAAC 180489
Oy 207 GluProTYrAlaAsnIleLYsAlaIleuGluAsp-----IleThrGlnIleuSerAla 224
Db 180488 GACGCC-----ATTAAAGAAACCTCGAATCGTTGGGCTGCGCAGTCCAAATAC 180438
Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 180437 TTTAATCAGGGGACACCTCAATACGACGACTCGCGGCTGAAAGAAATACCTCGGCGC 180378
Oy 237 -----ThAlaLeuValAlaAlaIleAlaValAlaGlyTYrTYrAsp 249
Db 180377 GCGAAACTCAATATCCAAATCAACGACGCAAGTAAACCAACTCGCGCCGACCGCTCAC 180318
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
Db 180317 ATCGCATCAGCATTTGACGAGGCGCAATCCGCAAAATCACCGACATGCAATTTGAA--- 180261
Oy 270 GlyIleuProValTYrIleAspTYrArgAlaValAlaGluValArgGlyGluGlyAlaAsp 289
Db 180260 GCGAACCAAGCTATTCGCGCGCAACTGATGCGG----- 180225
Oy 290 LysAlaPheThrThValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 180224 -----CAAAATGTCCTCGACGGAAGCGGCGCATTTGACATGG 180189
Oy 307 -----PheHisIleGlyLYsTYrGluThrLYsAsnLeuIleGlu 320
Db 180188 CTGACACGACGACCAATTCACAGCAGCAAAATTTGCCCAAGATATGGAAMAGTAAAC 180129
Oy 321 AsnAlaSerAlaGlnHisGlyTYrPheAspGlyArgTYrLeuAspArgSerValAspVal 340
Db 180128 GACTTCTACCAAAATTAACGGCTACTTCGATTCGTAATCCGTAACGACCAACCAAC 180069
Oy 341 IleuProAspAsnThAlaAspValSerIleuIleTYrAspThrGlyTYrThrGlnTYrArg 360
Db 180068 AACGAGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180009
Oy 361 PheAspGluValValPhePheThrIleAspProLYsThrAsnGlnIleuThrThrasPro 380
Db 180008 TCGGCGCAAGTC-----TCCATCGAAGCGGACCAACGAA----- 179973
Oy 381 AspLYsLeuProValLYsArgGluLeuLeuGluIleuLeuThrValAsnMetGlyGlu 400
Db 179972 -----GTCCCAAGCGCAACTGGAATAAAGTGTGACATGTAACCCGCGCAAA 179925
Oy 401 AlaTYrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTYrPhe 420
Db 179924 TCGTACGAACCGCAGCAGATACCGCGCTTTGGGT----- 179889
Oy 421 AsnMetValAsnThrGluIleValPheProGluArgGluIleGlnAsnAspGlnVal 440
Db 179888 -----GAGATTCAAGACCGCATGGCC 179868

```

Oy	441	SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu	460
Db	179867	TCGGCAGGCTACGCCATACAGC-----	GAATCAGCGTA 179835
Oy	461	GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--	479
Db	179834	CAGCGCGTGGCGAAGCGGTGAACCAAAACCGTCGATTTGGTCCGTCATCGAACCGGGC	179775
Oy	480	-----IleGluPheSerIleAspSerAsnLeuIleGlnAspGlyLeu	492
Db	179774	CGGAAATCTACGTCACAAAGATACATCCACGGCAACAAACAAACCCCGCAGAA--	179718
Oy	493	AsnLeuValAlaAlaIlysAlaArgHisIleuLeuYrAspMetProAsp-----	508
Db	179717	---GTCGTCGCCCGTGATTCAGCCAAATGGAATCCGCACTTACGACACTCCAACTG	179661
Oy	509	-----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle	523
Db	179660	CACGCTTCCAAAGACGGCGTGAGCTTTGGGCTACCTCCACATGCTCCAG-----	179610
Oy	524	LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleuProAspGluSer	543
Db	179609	---TTGATGCTCTCCCGCTTCCGCGCAGCCCGCAGAA--	179574
Oy	544	GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgGlySerProAla	563
Db	179573	---GTCGATTTG---AACATGATCTGACCGAACCTTCCACC-----	179538
Oy	564	AspValTYrGlnSerLysLysValProLeuTYrAlaPheValAlaSerAspLysProArg	583
Db	179538	-----	179538
Oy	584	AspGlyGlnIleGlyLeu---GlyTYrPcylSerAspThrGlyThrArgLeuValThr	601
Db	179537	---GGTCCCTCGATTTGAGCGGGGTGGGTGTCAGATACCGGG-----	179487
Oy	602	LysPheGlnHisAsnLeuIleAsnArgAspGlyTYrGlnAlaGlyAlaGluLeuArgLeu	621
Db	179486	TCCGCAGGCGTTTCCCAAGACACCTGTCGTCAGGGCAAGTCGGCGCAGCTGCGGCC	179427
Oy	622	SerGluAspLysGlyGlyValIleLysLeuTYrAlaThrLysProLeuSerHisProLeuAsn	641
Db	179426	TCCAGCAGCAAA-----ACCAGCGTTAACGGCTCGCGTGTACTGACCCGTACTTC	179373
Oy	642	AspGlnLeuArgAlaThrLeuGlyTYrGlnGlnGluValPheGlyHisSerThrAsnGly	661
Db	179372	ACGGCAGACGGGGTCAGCTGCGGTAC-----GATGTTTAAAGCAAGCCTTCGACCGC	179319
Oy	662	PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly	681
Db	179318	CGCAAGCATGACGACGACGATCAAAACATATTAACCCACACGGCA-----GGC	179271
Oy	682	GlyTYrAsnArg-----ThyTYrSerLeuArgTYrArgLeuAspLysLeu	696
Db	179270	GCAGCAGTCCGCATGAGCGGTGCTGTATTACGAATACGACGCGGTGATTTCCGGTTGGTG	179211
Oy	697	LysThrGlnAlaProProGluThrTYrProLinsLeuPro-----ValAspPheVal	713
Db	179210	GCAGCAACACTGACCGCTACACACTTACAAACAAAGCCGCAACACTGATGCGCACTTTATC	179151
Oy	714	Asn-----GlyLys---ProSerGlnGluValLeuAlaGlyValAlaValHisLys	730
Db	179150	AAGAATATCGGCAAAACCGAGCGCACACGGCAGCTTCAAAAGCTGCTACAAAGT	179091
Oy	731	ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTYrArg	744
Db	179090	ACCGTCGCTGGGGCGGCACAAACGACAGCCGCTTATGGCCGACGCGCGCTAC--	179034
Oy	745	GlnArgTYrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle	764
Db	179033	---CTGACGGGCGTGAACGCCGGAATC	179010

[illegible]





```

OY 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
      |||::: ||
Db 179210 GCAGAACACCTGACCGTACACCTACACAAAGCGCCCAACACTATGCCGACTTATTC 179151
      |||::: |||
OY 714 Asn-----GlyLys-----ProSerGlnGluAlaLeuGluAlaGlyValAlaValHisLys 730
      |||::: |||
Db 179150 AAGAAATACGGCAAAACCGACGACGACGAGCTTCAAAAGCTGCTGTACAAAGT 179091
      |||::: |||
OY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
      |||::: |||
Db 179090 ACCCTCGCGTGGGGCGGCAACAAACGACAGCCGTTATGGCCGACGCGGCTAC-- 179034
      |||::: |||
OY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
      |||::: |||
Db 179033 -----CTGACGGGCGGTGAACGCCGCAATC 179010
      |||::: |||
OY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
      |||::: |||
Db 179009 GCCCTGCGTGGCAGCAACACTGCATAC-----TACTCCGCC 178974
      |||::: |||
OY 785 AlaHisGlnMetThrGly----- 790
      |||::: |||
Db 178973 ACCCACAACCAACCTGTTCTCCCGCTGAGCAAAACCTTCACGCTGATGCTCGCGCC 178914
      |||::: |||
OY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
      |||::: |||
Db 178913 GAAGTCGGCATTCGCGGCGGTAC-----GGCAGAACCAAGCAAAATCCCTCTCTT 178863
      |||::: |||
OY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
      |||::: |||
Db 178862 GAAGACTTCTACGCGCGCGCGCTGCTGCTGCGGATACGAAAGCGGACGCTCGGT 178803
      |||::: |||
OY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGluValLeuAlaValGly----- 845
      |||::: |||
Db 178802 CCG-----AAAGTCTATGACGAATACGGCAAAATACAGCGCGGCAACAAA 178752
      |||::: |||
OY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
      |||::: |||
Db 178751 AAAGCCAAGTCTCCCGCCGACGCTCTTCCGATGCCCGCGCGAAGACGCGGCACCC 178692
      |||::: |||
OY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
      |||::: |||
Db 178691 GTCCGCGCTGAGCTGTTGCCGACGACGAGCGGTGGAGCGGCAAAACCTACGACGAC 178632
      |||::: |||
OY 870 ----- 870
      |||::: |||
Db 178631 AACAGCAGTTCCGCGCGACGCGGCAAGGTTCAACATTTACGCGCGCGCAATACCCAT 178572
      |||::: |||
OY 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrPheAsp 889
      |||::: |||
Db 178571 AAATCCACCTTTACCAACGATTCGCTATTCGCGCGGCGGCGGTATACSTGCTCTCG 178512
      |||::: |||
OY 890 ProValGlyGluValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
      |||::: |||
Db 178511 CCTTAGGCGCCATGAAATTCAGTACGCTACGCGCTGTAAGAAACCAACCGGAGACGCA 178452
      |||::: |||
OY 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
      |||::: |||
Db 178451 ATCCAAACGCTTCCATTCGCACTCGGCAAGAGCTTC 178416
      |||::: |||

```

Search completed: May 6, 2003, 03:37:45  
 Job time : 3534 secs

